# Chapter 6 Pricing of Exotic Options

In Chapter 4 we discussed the pricing of vanilla options (standard options) by means of finite differences. The methods were based on the simple partial differential equation (4.2),

$$\frac{\partial y}{\partial \tau} = \frac{\partial^2 y}{\partial x^2},$$

which was obtained from the Black–Scholes equation (4.1) for $V(S,t)$ via the transformations (4.3). These transformations have exploited the simple structure of the Black–Scholes operator and relied on the assumption of constant coefficients.

Exotic options lead to partial differential equations that are not of the simple structure of the basic Black–Scholes equation (4.1). In the general case, the transformations (4.3) are no longer useful and the PDEs must be solved directly. Thereby numerical instabilities or spurious solutions may occur that do not play any role for the methods of Chapter 4. To cope with the "new" difficulties, Chapter 6 introduces ideas and tools not needed in Chapter 4. Exotic options often involve higher-dimensional problems. This significantly adds to the complexity. An exhaustive discussion of the wide field of exotic options is beyond the scope of this book. The aim of this chapter will not be to formulate algorithms, but to give an outlook on several relevant aspects of computation, and on phenomena of stability. In this chapter, we still stick to the GBM model and move in the Black–Scholes world; for more general models see Chapter 7.

Sections 6.1 and 6.2 give a brief overview on important types of exotic options. Section 6.3 introduces approaches for path-dependent options, with the focus on Asian options. Then numerical aspects of convection-diffusion problems are discussed (in Section 6.4), and upwind schemes are analyzed (in Section 6.5). After these preparations, the Section 6.6 arrives at a state of the art high-resolution method. Finally, Section 6.7 will address penalty methods, with application to two-asset options.

R.U. Seydel, *Tools for Computational Finance*, Universitext,
DOI 10.1007/978-1-4471-2993-6_6, © Springer-Verlag London Limited 2012

## 6.1 Exotic Options

So far, this book has mainly concentrated on standard options. These are the American or European call or put options with vanilla payoff functions (1.1C) or (1.1P) as discussed in Section 1.1, based on a single underlying asset. The options traded on official exchanges are mainly standard options; there are market prices quoted in relevant newspapers.

All nonstandard options are called exotic options. That is, at least one of the features of a standard option is violated. One of the main possible differences between standard and exotic options lies in the payoff; examples are given in this section. Another extension from standard to exotic is an increase in the dimension, from single-factor to multifactor options; this will be discussed in Section 6.2. The distinctions between put and call, and between European and American options remain valid for exotic options.

Financial institutions have been imaginative in designing exotic options to meet the needs of clients. Many of the products have a highly complex structure. Exotic options are traded outside the exchanges (OTC), and often they are illiquid and no market prices are available. Then exotic options must be priced based on models. In general, their parameters are taken from the results obtained when standard options with comparable terms are calibrated to market prices. The simplest models extend the Black–Scholes model, which was summarized by Assumption 1.2.

Next we list some important types of exotic options. For more explanation we refer to [Hull00], [Wil98].

*Binary Option*: Binary options (or digital options) have a discontinuous payoff. For example, a binary put has the payoff

$$\Psi(S) := c \cdot \begin{cases} 1 & \text{if } S < K \\ 0 & \text{if } S \geq K \end{cases}$$

for a fixed amount $c$. See Figure 4.21 for an illustration of a binary call, and Section 3.5.5 for a two-dimensional example.

*Chooser Option*: After a specified period of time the holder of a chooser option can choose whether the option is a call or a put. The value of a chooser option at this time is

$$\max\{V_{\mathrm{C}}, V_{\mathrm{P}}\}$$

*Compound Option*: Compound options are options on options. Depending on whether the options are put or call, there are four main types of compound options. For example, the option may be a call on a call.

**Path-Dependent Options**

Options with payoff depending not only on the current value $S_T$ but also on the path of $S_t$ for previous times $t < T$ are called *path dependent*. Important

path-dependent options are the *barrier option*, the *lookback option*, and the *Asian option*.

*Barrier Option*: For a barrier option the payoff is contingent on the underlying asset's price $S_t$ reaching a certain threshold value $B$, which is called barrier. Barrier options can be classified depending on whether $S_t$ reaches $B$ from above *(down)* or from below *(up)*. Another feature of a barrier option is whether it ceases to exist when $B$ is reached *(knock out)*, or conversely comes into existence *(knock in)*. Obviously, for a down option, $S_0 > B$ and for an up option $S_0 < B$. Depending on whether the barrier option is a put or a call, several different types are possible. For example, the payoff of a European *down-and-out* call is

$$V_T = \begin{cases} (S_T - K)^+ & \text{in case } S_t > B \quad \text{for all } t \\ 0 & \text{in case } S_t \leq B \quad \text{for some } t \end{cases}$$

In the Black–Merton–Scholes framework, the value of the option before the barrier has been triggered still satisfies the Black–Scholes equation. The details of the barrier feature come in through the specification of boundary conditions [Wil98]. An example of an up-and-out call is illustrated in Figure 7.3, and a two-asset double barrier is discussed in Example 5.5.

*Lookback Option*: The payoff of a lookback option depends on the maximum or minimum value the asset price $S_t$ reaches during the life of the option. For example, the payoff of a lookback option is

$$\max_t S_t - S_T \, .$$

*Average Option / Asian Option*: The payoff from an Asian option depends on the average price of the underlying asset. This will be discussed in more detail in Section 6.3.

The exotic options of the above short list gain in complexity when they are multifactor options.

**Pricing of Exotic Options**

Several types of exotic options can be reduced to the Black–Scholes equation. In these cases the methods of Chapter 4 or Chapter 5 are adequate. In particular, barrier options under GBM are close to the standard options. For a knock-out option with barrier $B$, a boundary condition will be $V(B,t) = 0$, which is part of (4.19). Since their numerical treatment is widely analogous, we will not touch barrier options specifically.

For a number of options of the European type the Black–Scholes evaluation formula (A4.10) can be applied. For related reductions of exotic options we refer to [Hull00], [WiDH96], [Kwok98]. Approximations are possible with binomial methods or with Monte Carlo simulation. The Algorithm 3.6 applies, only the calculation of the payoff (step 2) must be adapted to the exotic option.

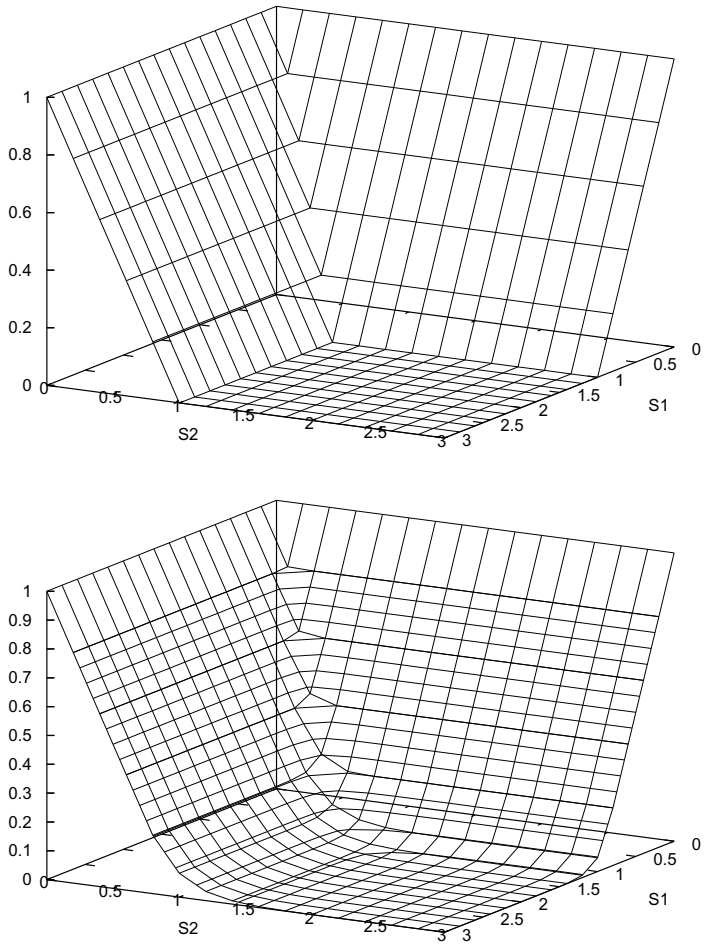


**Fig. 6.1.** Rainbow option of a put on the minimum of two assets; top: payoff $\Psi(S_1, S_2) = (1 - \min(S_1, S_2))^+$; bottom: $V(S_1, S_2, 0)$ approximated by a binomial method, level curves for slices with constant values of $S_1$, $S_2$, $V$

## 6.2 Options Depending on Several Assets

The options listed in Section 6.1 depend on one underlying asset. Options depending on several assets are discussed next. Two large groups of multifactor options are the *rainbow options* and the *baskets*. The subdivision into the groups is by their payoff. Assume $n$ underlying assets with prices $S_1, \ldots, S_n$. Different from the notation in previous chapters, the index refers to the number of the asset. Recall that two examples of exotic options with

two underlyings occurred earlier in this text: Example 3.8 of a binary put, and Section 5.4 with a basket-barrier call.

Rainbow options compare the value of individual assets [Smi97]. Examples of payoffs include

| | |
|---|---|
| $\max(S_1, \dots, S_n)$ | "$n$-color better-of option" |
| $\min(S_1, S_2)$ | "two-color worse-of option" |
| $(S_2 - S_1)^+$ | "outperformance option" |
| $(\min(S_1 - K, \dots, S_n - K))^+$ | "min call option" |
| $(S_2 - S_1 - K)^+$ | "spread call." |

Weights are possible too, for instance, $(c_1 S_2 - c_2 S_1)^+$. The outperformance option is also called spread option. Figure 6.1 (top) illustrates the payoff of a min put, and Figure 6.2 (bottom) the payoff of a max call. A basket is an option with payoff depending on a portfolio of assets. An example is the payoff of a basket call,

$$\left( \sum_{i=1}^{n} c_i S_i \, - K \right)^+ ,$$

where the weights $c_i$ are given by the portfolio. To gain a better feeling for such kind of options, it is recommendable to sketch the above payoffs for $n = 2$.

For the pricing of multifactor options the instruments introduced in the previous chapters apply. This holds for the four large classes of methods discussed before, namely, the PDE methods, the tree methods, the evaluation of integrals by quadrature, and the Monte Carlo methods. Each class subdivides into further methods.

For the choice of an appropriate method, the dimension $n$ is crucial. For large values of $n$, in particular PDE methods suffer from the curse of dimension ($\longrightarrow$ Exercise 4.18). At present state it is not possible to decide, above which threshold level of $n$ standard discretizations are too expensive.

**PDE methods** require relevant PDEs *and* boundary conditions. Often a Black–Merton–Scholes scenario is assumed. To extend the one-factor model, an appropriate generalization of geometric Brownian motion is needed. We begin with the two-factor model, with the prices of the two assets $S_1$ and $S_2$. The assumption of a constant-coefficient GBM is then expressed as

$$\begin{aligned} \mathrm{d}S_1 &= \mu_1 S_1 \, \mathrm{d}t + \sigma_1 S_1 \, \mathrm{d}W^{(1)} \\ \mathrm{d}S_2 &= \mu_2 S_2 \, \mathrm{d}t + \sigma_2 S_2 \, \mathrm{d}W^{(2)} \\ \mathsf{E}(\, \mathrm{d}W^{(1)} \, \mathrm{d}W^{(2)}) &= \rho \, \mathrm{d}t \,, \end{aligned} \tag{6.1a}$$

where $\rho$ is the correlation between the two assets, $-1 \le \rho \le 1$. Note that the third equation in (6.1a) is equivalent to $\mathsf{Cov}(\, \mathrm{d}W^{(1)}, \, \mathrm{d}W^{(2)}) = \rho \, \mathrm{d}t$, because $\mathsf{E}(\, \mathrm{d}W^{(1)}) = \mathsf{E}(\, \mathrm{d}W^{(2)}) = 0$. The correlation $\rho$ is given by the covariance of the returns $\frac{dS}{S}$, since

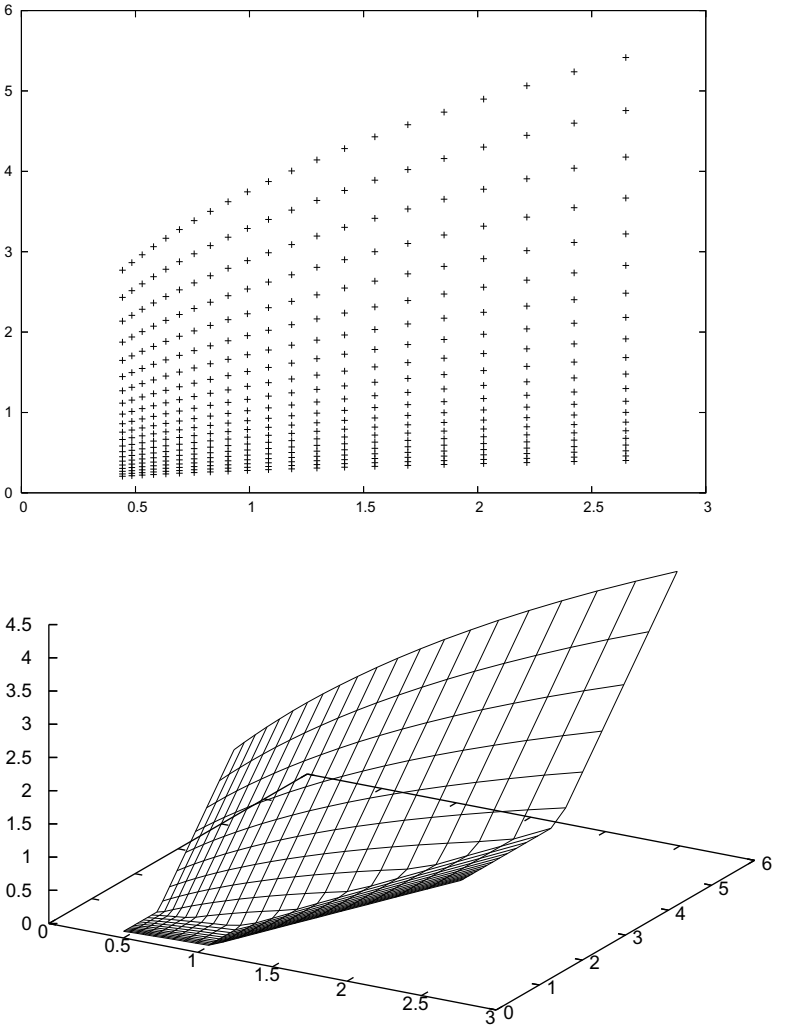


**Fig. 6.2.** Max call, with payoff $\Psi(S_1, S_2) = (\max(S_1, S_2) - K)^+$; numbers from Exercise 6.7; top: $(S_1, S_2)$-plane with the grid of the tree for the payoff, $t = T$, with $M = 20$; bottom: the payoff

$$\mathsf{Cov}\left(\frac{\mathrm{d}S_1}{S_1}, \frac{\mathrm{d}S_2}{S_2}\right) = \mathsf{E}(\sigma_1 \,\mathrm{d}W^{(1)} \sigma_2 \,\mathrm{d}W^{(2)}) = \rho\sigma_1\sigma_2 \,\mathrm{d}t\,. \tag{6.1b}$$

Compared to the more general system (1.41), the version (6.1a) with correlated Wiener processes has pulled out the scaling by the volatilities $\sigma_1, \sigma_2$. Then, following Section 2.3.3 and Exercise 2.9, the correlated Wiener processes can be decoupled by Cholesky decomposition of the correlation matrix

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} .$$

This leads to

$$\begin{aligned} \mathrm{d}W^{(1)} &= \mathrm{d}Z_1 \\ \mathrm{d}W^{(2)} &= \rho\,\mathrm{d}Z_1 + \sqrt{1-\rho^2}\,\mathrm{d}Z_2 \ , \end{aligned} \tag{6.1c}$$

where $Z_1$ and $Z_2$ are independent standard normally distributed processes. This was used already in (3.28). The resulting two-dimensional Black–Scholes equation was applied in Section 5.4, see equation (5.26). This is derived by the two-dimensional version of the Itô-Lemma ($\longrightarrow$ Appendix B2) and by a no-arbitrage argument. The resulting PDE (5.26) has independent variables $(S_1, S_2, t)$. Usually, the time variable is not counted when the dimension is discussed. In this sense, the PDE (5.26) is two-dimensional, whereas the classic Black–Scholes PDE (1.2) is considered as one-dimensional.

The general $n$-factor model is analogous. The appropriate GBM model is a straightforward generalization of (6.1a),

$$\begin{aligned} &\mathrm{d}S_i = (\mu_i - \delta_i)S_i\,\mathrm{d}t + \sigma_i S_i\,\mathrm{d}W^{(i)} \ , \quad i = 1, \ldots, n \\ &\mathsf{E}(\mathrm{d}W^{(i)}\mathrm{d}W^{(j)}) = \rho_{ij}\,\mathrm{d}t \ , \quad i, j = 1, \ldots, n \end{aligned} \tag{6.2a}$$

where $\rho_{ij}$ is the correlation between asset $i$ and asset $j$, and $\delta_i$ denotes a dividend flow rate paid by the $i$th asset. For a simulation of such a stochastic vector process see Section 2.3.3. The Black–Scholes-type PDE of the model (6.2a) is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{n} \rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} + \sum_{i=1}^{n}(r-\delta_i)S_i\frac{\partial V}{\partial S_i} - rV = 0\,. \tag{6.2b}$$

The derivation uses the general Itô formula (B2.1) ($\longrightarrow$ Exercise 6.5).

Boundary conditions depend on the specific type of option. For example in the "two-dimensional" situation in $(S_1, S_2, t)$-space, one boundary can be defined by the plane $S_1 = 0$ and the other by the plane $S_2 = 0$. It may be appropriate to apply the Black–Scholes vanilla formula (A4.10) along these planes, or to define one-dimensional sub-PDEs only for the purpose to calculate the values of $V(S_1, 0, t)$ and $V(0, S_2, t)$ along the boundary planes.

After the PDE with boundary conditions is set up, solutions are approximated by numerical methods. Standard discretizations are straightforward and work for small $n$. As a rule of thumb, for $n = 2$ and $n = 3$, such elementary PDE approaches are competitive to Monte Carlo. For large $n$, sparse-grid technology or multigrid are better choices, see the references in Section 3.5.1 and at the end of Chapter 4. Generally in a multidimensional situation, finite elements are recommendable. But FE methods suffer from the curse of dimension too. Irregular grids have been applied successfully [BeS08].

For **tree methods**, the binomial method can be generalized canonically [BoEG89]. (⟶ Exercise 6.7) But already for $n = 2$ the recombining standard tree with $M$ time levels requires $\frac{1}{3}M^3 + O(M^2)$ nodes, and for $n = 3$ the number of nodes is of the order $O(M^4)$. Tree methods also suffer from the curse of dimension. But obviously not all of the nodes of the canonical binomial approach are needed. The ultimate aim is to approximate the lognormal distribution, and this can be done with fewer nodes. Nodes in $\mathbb{R}^n$ should be constructed in such a way that the number of nodes grows comparably slower than the quality of the approximation of the distribution function. An example of a two-dimensional approach is presented in [Lyuu02]. Generalizing the trinomial approach to higher dimensions is not recommendable because of storage requirements, but other geometrical structures as icosahedral volumes can be applied. For different tree approaches, see [McW01]. For a convergence analysis of tree methods, and for an extension to Lévy processes, consult [FoVZ02], [MaSS06]. A tree approach that makes use of decoupling (similar as in Section 2.3.3) has shown to be favorable in multidimensional cases [KoM09].

An advantage of tree methods and of **Monte Carlo methods** is that no boundary conditions are needed. The essential advantage of MC methods is that they are much less affected by high dimensions, see the notes on Section 3.6. A correlation is achieved by $\mathrm{d}W = L\mathrm{d}Z$, where $LL^{tr}$ is the Cholesky decomposition of the $\rho$-matrix. An example of a five-dimensional American-style option is calculated in [BrG04], [LonS01], and one with dimension 30 in [Jon11]. It is most inspiring to perform Monte Carlo experiments on exotic options. For European-style options, this amounts to a straightforward application of Section 3.5 (⟶ Exercise 6.1).

## 6.3 Asian Options

The price of an Asian option[1] depends on the average price of the underlying and hence on the history of $S_t$. We choose this type of option to discuss some strategies of how to handle path-dependent options. Let us first define different types of Asian options via their payoff.

### 6.3.1 The Payoff

There are several ways how an average of past values of $S_t$ can be formed. If the price $S_t$ is observed at discrete time instances $t_i$, say equidistantly with time interval $h := T/n$, one obtains a times series $S_{t_1}, S_{t_2}, \ldots, S_{t_n}$. An obvious choice of average is the arithmetic mean

[1] Again, the name has no geographical relevance.

$$\frac{1}{n}\sum_{i=1}^{n} S_{t_i} = \frac{1}{T}\, h \sum_{i=1}^{n} S_{t_i}\,.$$

If we imagine the observation as continuously sampled in the time period $0 \le t \le T$, the above mean corresponds to the integral

$$\widehat{S} := \frac{1}{T}\int_0^T S_t \,\mathrm{d}t \tag{6.3}$$

The arithmetic average is used mostly. Sometimes the geometric average is applied, which can be expressed as

$$\left(\prod_{i=1}^{n} S_{t_i}\right)^{1/n} = \exp\left(\frac{1}{n}\log\prod_{i=1}^{n} S_{t_i}\right) = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\log S_{t_i}\right)\,.$$

Hence the continuously sampled geometric average of the price $S_t$ is the integral

$$\widehat{S}_{\mathrm{g}} := \exp\left(\frac{1}{T}\int_0^T \log S_t \,\mathrm{d}t\right)\,.$$

The averages $\widehat{S}$ and $\widehat{S}_{\mathrm{g}}$ are formulated for the time period $0 \le t \le T$, which corresponds to a European option. To allow for early exercise at time $t < T$, $\widehat{S}$ and $\widehat{S}_{\mathrm{g}}$ are modified appropriately, for instance to

$$\widehat{S} := \frac{1}{t}\int_0^t S_\theta \,\mathrm{d}\theta\,.$$

With an average value $\widehat{S}$ like the arithmetic average of (6.3) the payoff of Asian options can be written conveniently:

**Definition 6.1 (Asian option)**

With an average $\widehat{S}$ of the price evolution $S_t$ the payoff functions of Asian options are defined as

$$\begin{array}{ll} (\widehat{S}-K)^+ & \textit{average price call} \\ (K-\widehat{S})^+ & \textit{average price put} \\ (S_T-\widehat{S})^+ & \textit{average strike call} \\ (\widehat{S}-S_T)^+ & \textit{average strike put} \end{array}$$

The price options are also called *rate options*, or *fixed strike options*; the strike options are also called *floating strike options*. Compared to the vanilla payoffs of (1.1P), (1.1C), for an Asian price option the average $\widehat{S}$ replaces $S$ whereas for the Asian strike option $\widehat{S}$ replaces $K$. The payoffs of Definition 6.1 form surfaces on the quadrant $S > 0$, $\widehat{S} > 0$. The reader may visualize these payoff surfaces.

### 6.3.2 Modeling in the Black–Scholes Framework

The above averages can be expressed by means of the integral

$$A_t := \int_0^t f(S_\theta, \theta)\,\mathrm{d}\theta\,, \tag{6.4}$$

where the function $f(S,t)$ depends on the type of chosen average. In particular $f(S,t) = S$ corresponds to the continuous arithmetic average (6.3), up to scaling by the length of interval. For Asian options the price $V$ is a function of $S$, $A$ and $t$, which we write $V(S,A,t)$. To derive a partial differential equation for $V$ using a generalization of Itô's Lemma we require a differential equation for $A$. This is given by (6.4). Compare with (1.31) to see[2]

$$\begin{aligned} &\mathrm{d}A = a_A(t)\,\mathrm{d}t + b_A\,\mathrm{d}W_t\,, \\ &\text{with} \quad a_A(t) := f(S_t,t)\,, \quad b_A := 0\,. \end{aligned}$$

For $S_t$ the standard GBM of (1.33) is assumed. By the multidimensional version (B2.1) of Itô's Lemma adapted to $Y_t := V(S_t, A_t, t)$, the two terms in (1.44) or (1.45) that involve $b_A$ as factors to $\frac{\partial V}{\partial A}, \frac{\partial^2 V}{\partial A^2}$ vanish. Accordingly,

$$\mathrm{d}V_t = \left(\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + f(S,t)\frac{\partial V}{\partial A}\right)\mathrm{d}t + \sigma S\frac{\partial V}{\partial S}\,\mathrm{d}W_t\,.$$

The derivation of the Black–Scholes-type PDE goes analogously as outlined in Appendix A4 for standard options and results in

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} + f(S,t)\frac{\partial V}{\partial A} - rV = 0\,. \tag{6.5}$$

Compared to the original vanilla version (1.2), only one term in (6.5) is new, namely,

$$f(S,t)\frac{\partial V}{\partial A}\,.$$

As we will see below, the lack of a second-order derivative with respect to $A$ may cause numerical difficulties. The transformations (4.3) cannot be applied advantageously to (6.5). — As an alternative to the definition of $A_t$ in (6.4), one can scale by $t$. This leads to a different "new term" (⟶ Exercise 6.2e).

[2] The ordinary integral $A_t$ is random but has zero quadratic variation [Shr04].

### 6.3.3 Reduction to a One-Dimensional Equation

Solutions to (6.5) are defined on the domain

$$S > 0\ ,\ A > 0\ ,\ 0 \le t \le T$$

of the $(S, A, t)$-space. The extra $A$-dimension leads to significantly higher costs when (6.5) is solved numerically. This is the general situation. But in some cases it is possible to reduce the dimension. Let us discuss an example, concentrating on the case $f(S,t) = S$ of the arithmetic average.

We consider a European arithmetic average strike (floating strike) call with payoff

$$\left(S_T - \frac{1}{T}A_T\right)^+ = S_T\left(1 - \frac{1}{TS_T}\int_0^T S_\theta \,\mathrm{d}\theta\right)^+ .$$

An auxiliary variable $R_t$ is defined by

$$R_t := \frac{1}{S_t}\int_0^t S_\theta \,\mathrm{d}\theta\,,$$

and the payoff is rewritten

$$S_T\left(1 - \frac{1}{T}R_T\right)^+ = S_T \cdot \text{function}(R_T, T)\,. \tag{6.6}$$

This motivates trying a separation of the solution in the form

$$V(S, A, t) = S \cdot H(R, t) \tag{6.7}$$

for some function $H(R,t)$. In this role, $R$ is an independent variable. From (6.6) the payoff follows:

$$H(R_T, T) = (1 - \tfrac{1}{T}R_T)^+ \tag{6.8a}$$

Substituting the separation ansatz (6.7) into the PDE (6.5) leads to a PDE for $H$,

$$\frac{\partial H}{\partial t} + \frac{1}{2}\sigma^2 R^2 \frac{\partial^2 H}{\partial R^2} + (1 - rR)\frac{\partial H}{\partial R} = 0 \tag{6.8b}$$

($\longrightarrow$ Exercise 6.2c). To solve this PDE, boundary conditions are required. Their choice in general is not unique. The following considerations from [WiDH96] suggest boundary conditions.

A right-hand boundary condition for $R \to \infty$ follows from the payoff (6.8a), which implies $H(R_T, T) = 0$ for $R_T \to \infty$. The integral $A_t = S_t R_t$ is bounded, hence $S \to 0$ for $R \to \infty$. For $S \to 0$ a European call option is not exercised, which suggests to prescribe the boundary condition

$$H(R, t) = 0 \quad \text{for} \quad R \to \infty \text{ and all } t\,. \tag{6.9}$$

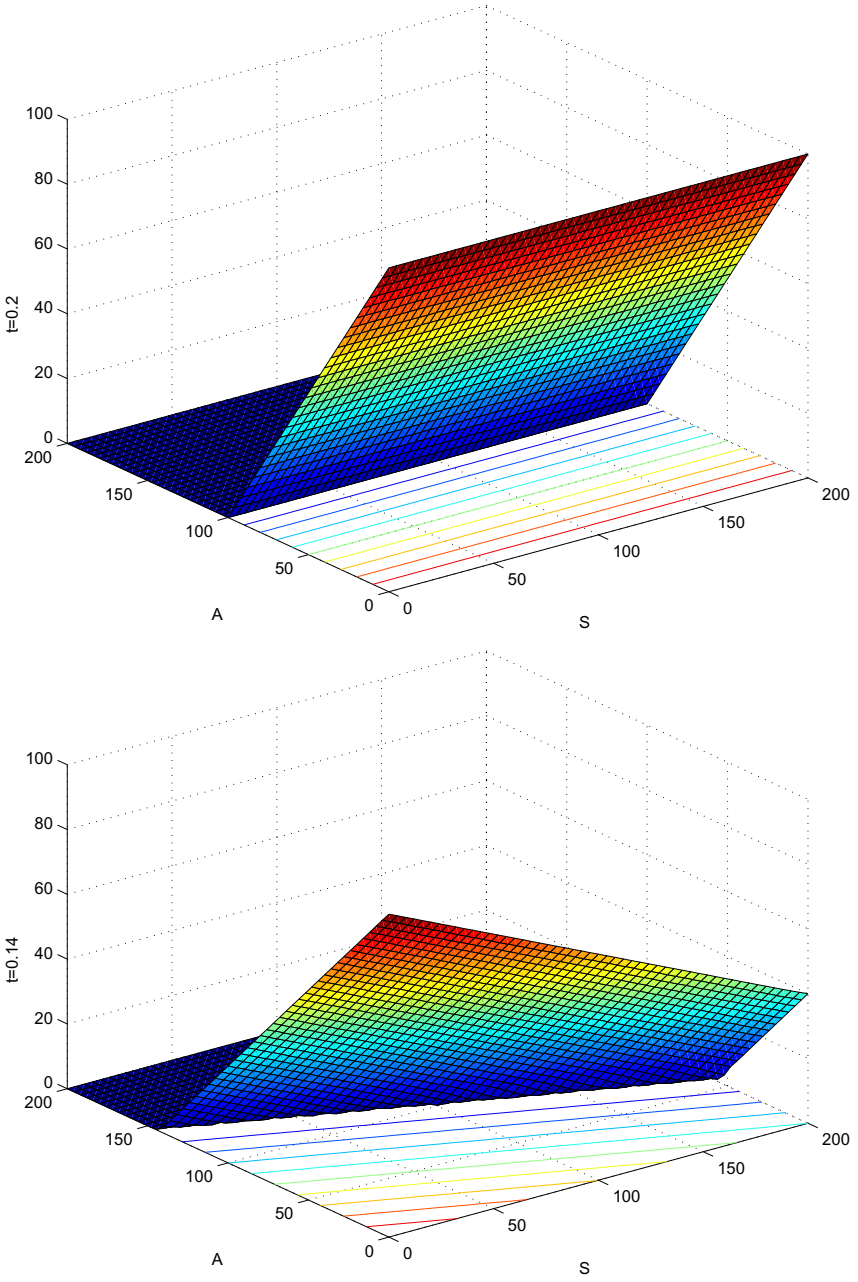


**Fig. 6.3.** Asian European fixed strike put, $K = 100$, $T = 0.2$, $r = 0.05$, $\sigma = 0.25$, payoff ($t = 0.2$) and three solution surfaces for $t = 0.14$, $t = 0.06$, and $t = 0$. With kind permission of Sebastian Göbel. (Figure continued on facing page)

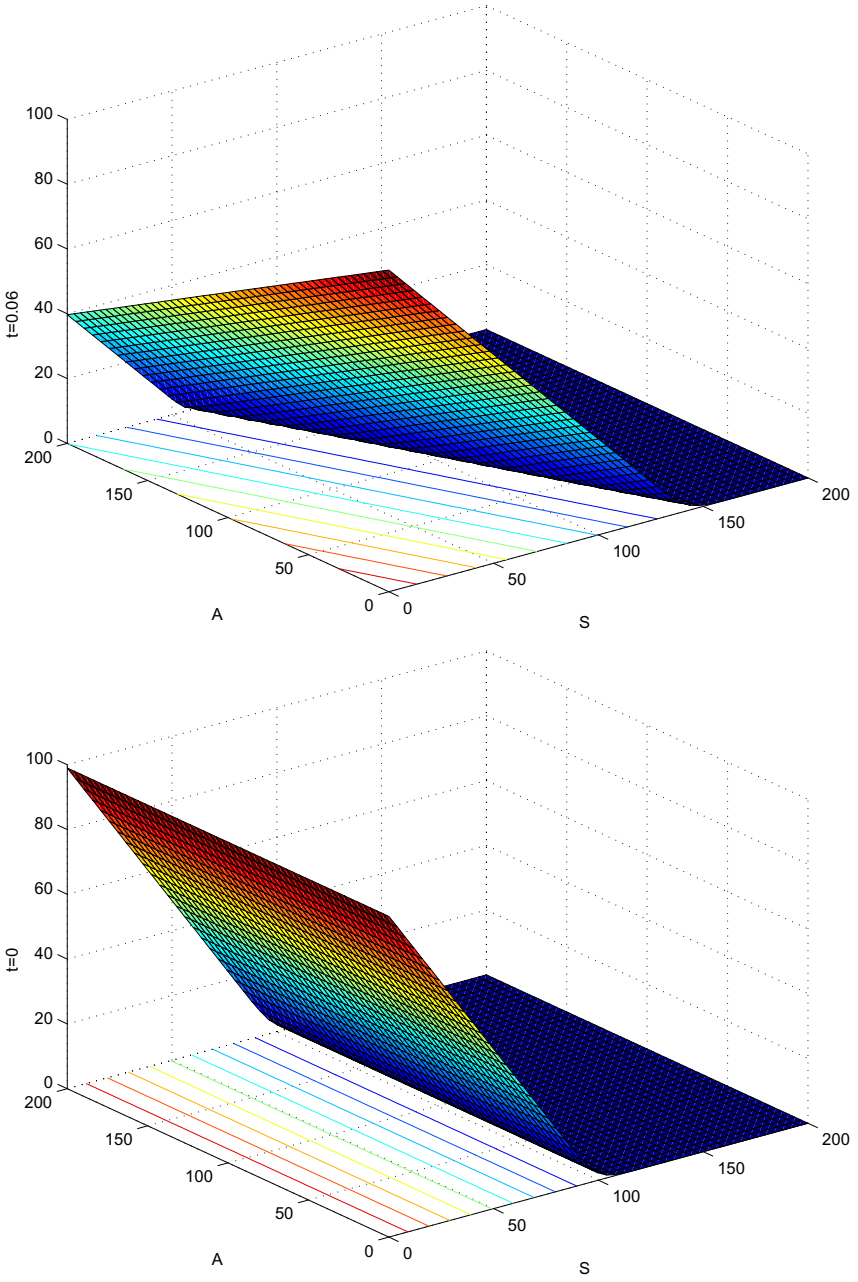


**Fig. 6.3.** continued

At the left-hand boundary $R = 0$ we encounter more difficulties. Note that the integral $R_t$ satisfies the SDE

$$\mathrm{d}R_t = (1 + (\sigma^2 - \mu)R_t)\,\mathrm{d}t - \sigma R_t\,\mathrm{d}W_t$$

($\longrightarrow$ Exercise 6.2d). Even if $R_0 = 0$ holds, this SDE shows that $\mathrm{d}R_0 = \mathrm{d}t$ and $R_t$ will not stay at 0. So there is no reason to expect $R_T = 0$, and the value of the payoff cannot be predicted. Another kind of boundary condition is required.

To this end, we start from the PDE (6.8b), which for $R \to 0$ is equivalent to

$$\frac{\partial H}{\partial t} + \frac{1}{2}\sigma^2 R^2 \frac{\partial^2 H}{\partial R^2} + \frac{\partial H}{\partial R} = 0\,.$$

Assuming that $H$ is bounded, one can prove that the term

$$R^2 \frac{\partial^2 H}{\partial R^2}$$

vanishes for $R \to 0$. The resulting boundary condition is

$$\frac{\partial H}{\partial t} + \frac{\partial H}{\partial R} = 0 \quad \text{for} \quad R \to 0\,. \tag{6.10}$$

The vanishing of the second-order derivative term is shown by contradiction: Assuming a nonzero value of $R^2 \frac{\partial^2 H}{\partial R^2}$ leads to

$$\frac{\partial^2 H}{\partial R^2} = O\left(\frac{1}{R^2}\right)\,,$$

which can be integrated twice to

$$H = O(\log R) + c_1 R + c_2\,.$$

This contradicts the boundedness of $H$ for $R \to 0$.

For a numerical realization of the boundary condition (6.10) in the finite-difference framework of Chapter 4, we may use the second-order formula

$$\left.\frac{\partial H}{\partial R}\right|_{0,\nu} = \frac{-3H_{0,\nu} + 4H_{1,\nu} - H_{2,\nu}}{2\Delta R} + O(\Delta R^2)\,. \tag{6.11}$$

The indices have the same meaning as in Chapter 4. We summarize the boundary-value problem of PDEs in (6.12).

$$\begin{aligned}
&\frac{\partial H}{\partial t} + \frac{1}{2}\sigma^2 R^2 \frac{\partial^2 H}{\partial R^2} + (1 - rR)\frac{\partial H}{\partial R} = 0\\
&H(R_T, T) = \left(1 - \tfrac{R_T}{T}\right)^+\\
&H = 0 \quad \text{for} \quad R \to \infty\\
&\frac{\partial H}{\partial t} + \frac{\partial H}{\partial R} = 0 \quad \text{for} \quad R = 0
\end{aligned} \tag{6.12}$$

Solving this problem numerically for $0 \le t \le T$, $R \ge 0$, gives $H(R,t)$, and via (6.7) the required values of $V$.

### 6.3.4 Discrete Monitoring

Instead of defining a continuous averaging as in (6.3), a realistic scenario is to assume that the average is monitored only at discrete time instances

$$t_1, t_2, \ldots, t_M\,.$$

These time instances are not to be confused with the grid times of the numerical discretization. The discretely sampled arithmetic average at $t_k$ is given by

$$A_{t_k} := \frac{1}{k}\sum_{i=1}^{k} S_{t_i},\ k = 1, \ldots, M\,. \tag{6.13}$$

A new average is updated from the previous one by

$$A_{t_k} = A_{t_{k-1}} + \frac{1}{k}(S_{t_k} - A_{t_{k-1}})$$

or

$$A_{t_{k-1}} = A_{t_k} + \frac{1}{k-1}(A_{t_k} - S_{t_k})\,.$$

The latter of these update formulas is relevant to us, because we integrate backwards in time. The discretely sampled $A_t$ is constant between consecutive sampling times, and $A$ jumps at $t_k$ with the step

$$\frac{1}{k-1}(A_{t_k} - S_{t_k})\,.$$

For each $k$ this jump can be written

$$A^-(S) = A^+(S) + \frac{1}{k-1}(A^+(S) - S),\ \text{where } S = S_{t_k}\,. \tag{6.14a}$$

$A^-$ and $A^+$ denote the values of $A$ immediately before and immediately after sampling at $t_k$. The no-arbitrage principle implies continuity of $V$ at the sampling instances $t_k$ in the sense of continuity of $V(S_t, A_t, t)$ for any realization of a random walk. In our setting, this continuity is written

$$V(S, A^+, t_k) = V(S, A^-, t_k)\,. \tag{6.14b}$$

But for a *fixed* $(S, A)$ the equations (6.14a/b) define a **jump** of $V$ at $t_k$.

The numerical application of the jump condition (6.14) is as follows: The $A$-axis is discretized into discrete values $A_j$, $j = 1, \ldots, J$. For each time period between two consecutive sampling instances, say for $t_{k+1} \to t_k$, the option's value is independent of $A$ because in our discretized setting $A_t$ is piecewise constant; accordingly $\frac{\partial V}{\partial A} = 0$ in (6.5). Based on this semi-discretization, $J$

one-dimensional Black–Scholes equations are integrated separately and independently for the short time interval from $t_{k+1}$ to $t_k$, one BS-equation for each $j$. Each of the one-dimensional Black–Scholes problems has its own "terminal" condition to start from. For each $A_j$, the "first" terminal condition for $t_M = T$ is taken from the payoff surface. Proceeding backwards in time, at each sampling time $t_k$ the $J$ parallel one-dimensional Black–Scholes problems are halted because new terminal conditions must be derived from the jump condition (6.14). The new values for $V(S, A_j, t_k)$ that serve as terminal values (starting values for the backward integration) for the next time period $t_k \to t_{k-1}$, are defined by the jump condition. Since $A_j + \frac{1}{k-1}(A_j - S)$ in general does not agree with one of the node values $A_j$, interpolation is applied. Hence the starting function for the next BS-step for $A = A_j$ can be written

$$V^{\text{interpol}}(S,\, A + \frac{1}{k-1}(A - S),\, t_k)\,.$$

Only at these sampling times $t_k$ the $J$ standard one-dimensional Black–Scholes problems are coupled; the coupling is provided by the interpolation. In this way, a sequence of surfaces $V(S, A, t_k)$ is approximated for $t_M = T, \ldots, t_1 = 0$ in a line-wise fashion. Figure 6.3 shows[3] the payoff and three surfaces calculated for an Asian European fixed strike put. As this illustration indicates, there is a kind of rotation of this surface as $t$ varies from $T$ to 0.

## 6.4 Numerical Aspects

A direct numerical approach to the PDE (6.5) for functions $V(S, A, t)$ depending on three independent variables requires more effort than in the $V(S, t)$-case. For example, a finite-difference approach uses a three-dimensional grid. And a separation ansatz as in Section 5.3 applies with two-dimensional basis functions. Although much of the required technology is widely analogous to the approaches discussed in Chapters 4 and 5, a thorough numerical treatment of higher-dimensional PDEs is beyond the scope of this book. Here we confine ourselves to PDEs with two independent variables, as in (6.8b).

### 6.4.1 Convection-Diffusion Problems

Before entering a discussion on how to solve numerically a PDE like (6.8b) without using transformations like (4.3), we perform an experiment with our well-known "classical" Black–Scholes equation (1.2). In contrast to the procedure of Chapter 4 we directly apply finite-difference quotients to (1.2). Here we use the second-order differences of Section 4.2.1 for a European call,

[3] after interpolation; MATLAB graphics; similar [ZvFV99]

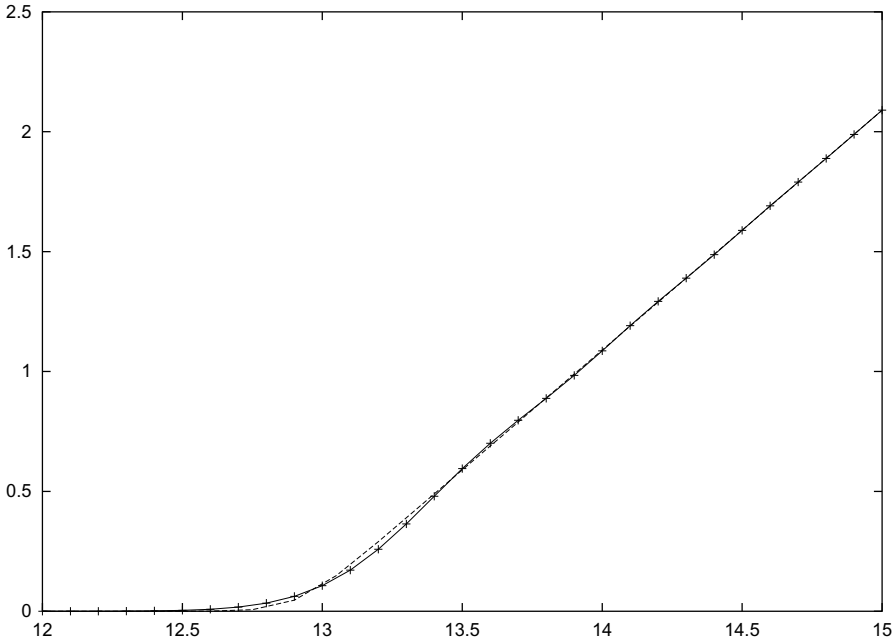


**Fig. 6.4.** European call, $K = 13$, $r = 0.15$, $\sigma = 0.01$, $T = 1$. Crank–Nicolson approximation $V(S, 0)$ with $\Delta t = 0.01$, $\Delta S = 0.1$ and centered difference scheme for $\frac{\partial V}{\partial S}$. Comparison with the exact Black–Scholes values (dashed).

and compare the numerical approximation with the exact solution (A4.10). Figure 6.4 shows the result for $V(S, 0)$. The lower part of the figure depicts an oscillating error, which seems to be small. But differentiating magnifies oscillations. This is clearly visible in Figure 6.5, where the important hedge variable delta= $\frac{\partial V}{\partial S}$ is depicted. The wiggles are even worse for the second-order derivative gamma. These oscillations are financially unrealistic and are not tolerable, and we have to find its causes. The oscillations are *spurious* in that they are produced by the numerical scheme and are not solutions of the differential equation. The spurious oscillations do not exist for the transformed version $y_\tau = y_{xx}$, which is illustrated by Figure 6.6.

In order to understand possible reasons why spurious oscillations may occur, we invoke elementary fluid dynamics, where so-called convection-diffusion equations play an important role. For such equations, the second-order term is responsible for diffusion and the first-order term for convection. The ratio of convection to diffusion (their coefficients, scaled by a characteristic length) is the *Péclet number*, a dimensionless parameter characterizing the convection-diffusion problem. It turns out that the Péclet number is relevant for the understanding of underlying phenomena. Let us see what the Péclet numbers are for several PDEs discussed so far in the text.

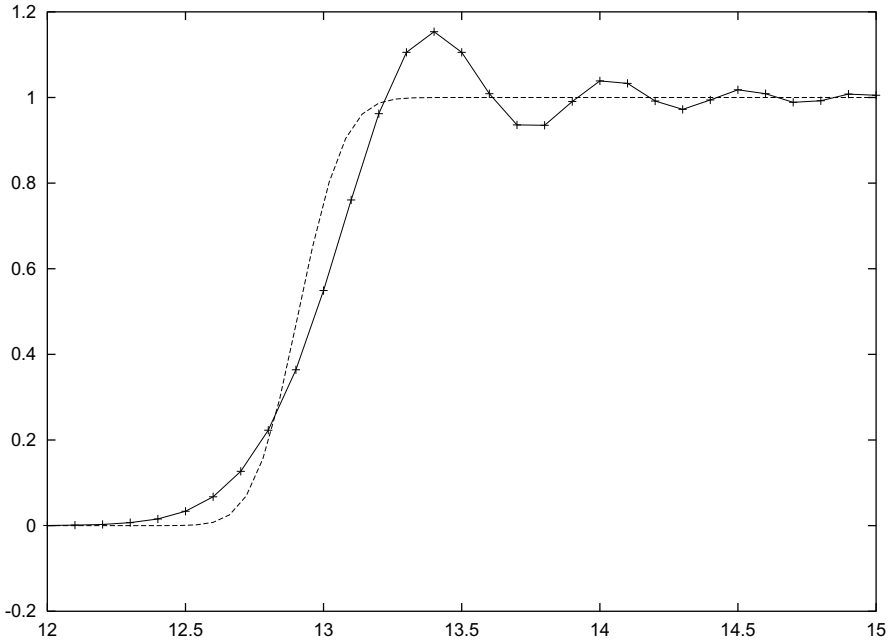


**Fig. 6.5.** Delta= $\frac{\partial V}{\partial S}$, otherwise the same data as in Figure 6.4

As a first example we take the original Black–Scholes equation (1.2), with

$$\begin{aligned} &\text{diffusion term:} && \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \\ &\text{convection term:} && rS\frac{\partial V}{\partial S} \\ &\text{length scale:} && \Delta S \end{aligned}$$

When the coefficients —not the derivatives— enter the Péclet number, and $\Delta S$ is taken as characteristic length, the number is

$$\frac{rS}{\frac{1}{2}\sigma^2 S^2}\Delta S = \frac{2r}{\sigma^2}\ \frac{\Delta S}{S}\,.$$

Since this dimensionless parameter involves the mesh size $\Delta S$ it is also called *mesh Péclet number*.[4] Experimental evidence indicates that the higher the Péclet number, the higher the danger that the numerical solution exhibits oscillations.

Next we examine other PDEs for their Péclet numbers: The PDE $y_\tau = y_{xx}$ has no convection term, hence its Péclet number is zero. Asian options

[4] In case of a continuous dividend flow $\delta$, replace $r$ by $r - \delta$.

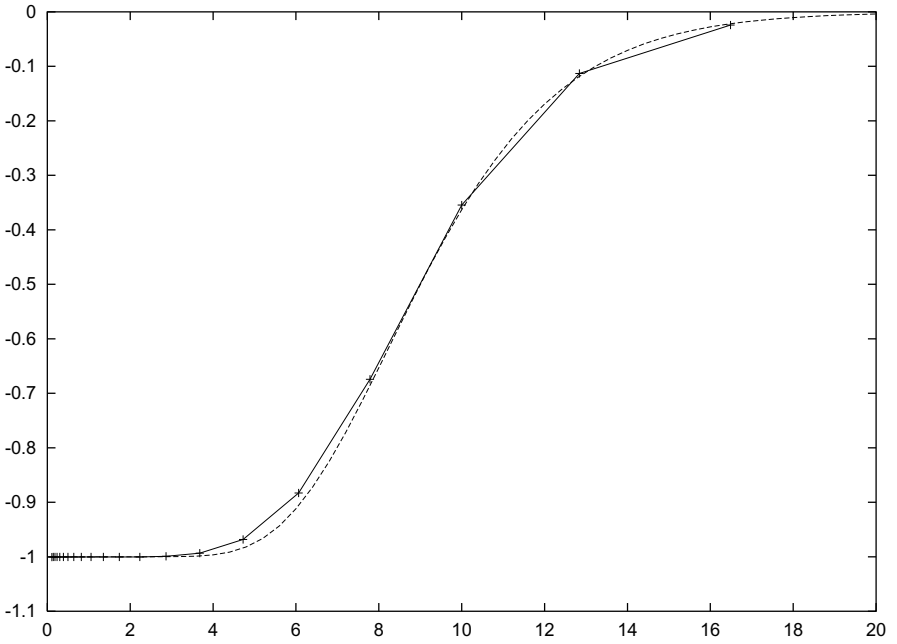


**Fig. 6.6.** European put, $K = 10$, $r = 0.06$, $\sigma = 0.30$, $T = 1$. Approximation delta$= \frac{\partial V}{\partial S}(S, 0)$ based on $y_\tau = y_{xx}$ with $m = 40$. Comparison with the exact Black–Scholes values (dashed).

described by the PDE (6.5) have a cumbersome situation: With respect to $A$ there is no diffusion term (i.e., no second-order derivative), hence its Péclet number is $\infty$! For the original Black–Scholes equation the Péclet number basically amounts to $r/\sigma^2$. It may become large when a small volatility $\sigma$ is not compensated by a small riskless interest rate $r$. And for the reduced PDE (6.8b), the Péclet number is

$$\frac{\Delta R(1 - rR)}{\frac{1}{2}\sigma^2 R^2} ,$$

here a small $\sigma$ can not be compensated by a small $r$.

These investigations of the Péclet numbers do not yet explain *why* spurious oscillations occur, but should open our eyes to the relation between convection and diffusion in the different PDEs. Let us discuss causes of the oscillations by means of a **model problem**. The model problem is the pure initial-value problem for a scalar function $u$ defined on $t \geq 0$, $x \in \mathbb{R}$,

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = b\frac{\partial^2 u}{\partial x^2} , \quad u(x, 0) = u_0(x) . \tag{6.15}$$

We assume $b \geq 0$. This sign of $b$ does not contradict the signs in (6.8b) since there we have a terminal condition for $t = T$, whereas (6.15) prescribes an

initial condition for $t = 0$. The equation (6.15) is meant to be integrated in forward time with discretization step size $\Delta t > 0$. So the equation (6.15) is a model problem representing a large class of convection-diffusion problems, to which the equation (6.8b) belongs. For the Black–Scholes equation, the simple transformation $S = K\mathrm{e}^x$, $t = T - \tau$, which works even for variable coefficients $r, \sigma$, produces (6.15) except for a further term $-ru$ on the right-hand side (compare Exercise 1.2). And for constant $r, \sigma$ the transformed equation $y_\tau = y_{xx}$ is a member of the class (6.15), although it lacks convection. Discussing the stability properties of the model problem (6.15) will help us understanding how discretizations of (1.2) or (6.8b) behave. For the analysis assume an equidistant grid on the $x$-range, with grid size $\Delta x > 0$ and nodes $x_j = j\Delta x$ for integers $j$. And for sake of simplicity, assume $a$ and $b$ are constants.

### 6.4.2 Von Neumann Stability Analysis

First we apply to (6.15) the standard second-order centered space difference schemes in $x$-direction together with a forward time step, leading to

$$\frac{w_{j,\nu+1} - w_{j,\nu}}{\Delta t} + a\frac{w_{j+1,\nu} - w_{j-1,\nu}}{2\Delta x} = b\delta_{xx} w_{j,\nu} \tag{6.16}$$

with $\delta_{xx} w_{j,\nu}$ defined as in (4.13). This scheme is called *Forward Time Centered Space* (FTCS). "Forward time" reflects the explicit (forward) Euler step, and "centered space" refers to our well-established second-order difference quotients. Instead of performing an eigenvalue-based stability analysis as in Chapter 4, we now apply the von Neumann stability analysis. This method expresses the approximations $w_{j,\nu}$ of the $\nu$-th time level by a sum of *eigenmodes* or Fourier modes,

$$w_{j,\nu} = \sum_k c_k^{(\nu)} \mathrm{e}^{\mathrm{i}k\eta j\Delta x}\,, \tag{6.17}$$

where i denotes the imaginary unit, and $k\eta$ are the *wave numbers* with fundamental wave number[5] $\eta := 2\pi/L$. A set of coefficients $c_k^{(\nu)}$ in (6.17) exists for each time level $t_\nu$, it is the basis of the discrete Fourier transform (C1.8), which takes numbers $w_j$ into coefficients $c_k$, and back. Substituting the expression (6.17) into the FTCS-difference scheme (6.16) leads to a corresponding sum for $w_{j,\nu+1}$ with coefficients $c_k^{(\nu+1)}$ ($\longrightarrow$ Exercise 6.3). The linearity of the scheme (6.16) allows to find a relation

$$c_k^{(\nu+1)} = G_k c_k^{(\nu)}\,,$$

[5] $L$ stands for the wave length or the length of the interval. In case of a partition into $n$ steps of size $\Delta x$, $\eta\Delta x = 2\pi/n$. Since $\eta$ will drop out below, we may set $\eta = 1$ for the following analysis. It will be sufficient to study the propagation of $\mathrm{e}^{\mathrm{i}kx}$.

where $G_k$ is the *growth factor* of the mode with wave number $k$. In case $|G_k| \le 1$ holds, it is guaranteed that the modes $\mathrm{e}^{\mathrm{i}kx}$ in (6.17) are not amplified, which means the method is stable. This parallels Lemma 4.2 without the need of calculating eigenvalues.

Applying the von Neumann stability analysis to (6.16) leads to

$$G_k = 1 - 2\lambda + \left(\tfrac{\gamma}{2} + \lambda\right) \mathrm{e}^{-\mathrm{i}k\eta\Delta x} + \left(\lambda - \tfrac{\gamma}{2}\right) \mathrm{e}^{\mathrm{i}k\eta\Delta x} ,$$

where we use the abbreviations

$$\gamma := \frac{a\Delta t}{\Delta x} \quad , \quad \lambda := \frac{b\Delta t}{\Delta x^2} \quad , \quad \beta := \frac{a\Delta x}{b} . \tag{6.18}$$

Here $\gamma = \beta\lambda$ is the famous *Courant number*, and $\beta$ is the mesh Péclet number. For a finite value of the latter, assume $b > 0$. Using $\mathrm{e}^{\mathrm{i}\alpha} = \cos\alpha + \mathrm{i}\sin\alpha$ and

$$s := \sin\frac{k\eta\Delta x}{2} , \quad \cos k\eta\Delta x = 1 - 2s^2 , \quad \sin k\eta\Delta x = 2s\sqrt{1-s^2} ,$$

we arrive at

$$G_k = 1 - 2\lambda + 2\lambda\cos k\eta\Delta x - \mathrm{i}\beta\lambda\sin k\eta\Delta x \tag{6.19}$$

and

$$|G_k|^2 = (1 - 4\lambda s^2)^2 + 4\beta^2\lambda^2 s^2(1 - s^2) .$$

A straightforward discussion of this polynomial on $0 \le s^2 \le 1$ reveals that $|G_k| \le 1$ for

$$0 \le \lambda \le \tfrac{1}{2} \quad , \quad \lambda\beta^2 \le 2 . \tag{6.20}$$

The inequality $0 \le \lambda \le \frac{1}{2}$ brings back the stability criterion of Section 4.2.4. The inequality $\lambda\beta^2 \le 2$ is an additional restriction to the parameters $\lambda$ and $\beta$. Because of

$$\lambda\beta^2 = \frac{a^2\Delta t}{b}$$

this restriction depends on the discretization steps $\Delta t$, $\Delta x$, and on the convection parameter $a$ and the diffusion parameter $b$ as defined in (6.18). The restriction due to the convection becomes apparent when we, for example, choose $\lambda = \frac{1}{2}$ for a maximal time step $\Delta t$. Then $|\beta| \le 2$ is a bound imposed on the mesh Péclet number, which restricts $\Delta x$ to $\Delta x \le 2b/|a|$. A violation of this bound might be an explanation why the difference schemes of (6.16) applied to the Black–Scholes equation (1.2) exhibit faulty oscillations.[6] The bounds on $|\beta|$ and $\Delta x$ are not active for problems without convection ($a = 0$). Note that the bounds give a severe restriction on problems with small values of the diffusion constant $b$. For $b \to 0$ (no diffusion) and $a \neq 0$ we encounter the consequence $\Delta t \to 0$, and the scheme (6.16) can not be applied at all. Although the constant-coefficient model problem (6.15) is not the same

[6] In fact, the situation is more subtle. We postpone an outline of how *dispersion* is responsible for the oscillations to Section 6.5.2.

as the Black–Scholes equation (1.2) or the equation (6.8b), the above analysis reflects the core of the difficulties. We emphasize that small values of the volatility represent small diffusion. So other methods than the standard finite-difference approach (6.16) are needed.

## 6.5 Upwind Schemes and Other Methods

The instability analyzed for the model combination (6.15)/(6.16) occurs when the mesh Péclet number is high and because the symmetric and centered difference quotient is applied to the first-order derivative. Next we discuss the extreme case of an infinite Péclet number of the model problem, namely, $b = 0$. The resulting PDE is the prototypical equation

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0\,. \tag{6.21}$$

### 6.5.1 Upwind Scheme

The standard FTCS approach for (6.21) does not lead to a stable scheme. The PDE (6.21) has solutions in the form of *traveling waves*,

$$u(x,t) = F(x - at)\,,$$

where $F(\xi) = u_0(\xi)$ in case initial conditions $u(x,0) = u_0(x)$ are incorporated. For $a > 0$, the profile $F(\xi)$ drifts in positive $x$-direction: the "wind blows to the right." Seen from a grid point $(j,\nu)$, the neighboring node $(j-1,\nu)$ lies *upwind* and $(j+1,\nu)$ lies *downwind*. Here the $j$ indicates the node $x_j$ and $\nu$ the time instant $t_\nu$. Information flows from upstream to downstream nodes. Accordingly, the first-order difference scheme

$$\frac{w_{j,\nu+1} - w_{j,\nu}}{\Delta t} + a\frac{w_{j,\nu} - w_{j-1,\nu}}{\Delta x} = 0 \tag{6.22}$$

is called *upwind discretization* ($a > 0$). The scheme (6.22) is also called Forward Time Backward Space (FTBS) scheme.

Applying the von Neumann stability analysis to the scheme (6.22) leads to growth factors given by

$$G_k := 1 - \gamma + \gamma \mathrm{e}^{-\mathrm{i}k\eta\Delta x}\,. \tag{6.23}$$

Here $\gamma = \frac{a\Delta t}{\Delta x}$ is the Courant number from (6.18). As in Subsection 6.4.2, the stability requirement $|G_k| \le 1$ should hold such that the coefficients $c_k^{(\nu)}$ remain bounded for all $k$ and $\nu \to \infty$. It is easy to see that

$$\gamma \le 1 \;\Rightarrow\; |G_k| \le 1\,.$$

(The reader may sketch the complex $G$-plane to realize the situation.) The condition $|\gamma| \leq 1$ is called the **Courant–Friedrichs–Lewy (CFL) condition**. The above analysis shows that this condition is sufficient to ensure stability of the upwind-scheme (6.22) applied to the PDE (6.21) with prescribed initial conditions.

In case $a < 0$, the scheme in (6.22) is no longer an upwind scheme. The upwind scheme for $a < 0$ is

$$\frac{w_{j,\nu+1} - w_{j,\nu}}{\Delta t} + a\frac{w_{j+1,\nu} - w_{j,\nu}}{\Delta x} = 0 \tag{6.24}$$

The von Neumann stability analysis leads to the restriction $|\gamma| \leq 1$, or $\lambda|\beta| \leq 1$ if expressed in terms of the mesh Péclet number, see (6.18). This again emphasizes the importance of small Péclet numbers.

We note in passing that the FTCS scheme for $u_t + au_x = 0$, which is unstable, can be cured by replacing $w_{j,\nu}$ by the average of its two neighbors. The resulting scheme

$$w_{j,\nu+1} = \tfrac{1}{2}(w_{j+1,\nu} + w_{j-1,\nu}) - \tfrac{1}{2}\gamma(w_{j+1,\nu} - w_{j-1,\nu}) \tag{6.25}$$

is called *Lax–Friedrichs scheme*. It is stable if and only if the CFL condition is satisfied. A simple calculation shows that the Lax–Friedrichs scheme (6.25) can be rewritten in the form

$$\frac{w_{j,\nu+1} - w_{j,\nu}}{\Delta t} = -a\frac{w_{j+1,\nu} - w_{j-1,\nu}}{2\Delta x} + \frac{1}{2\Delta t}\left(w_{j+1,\nu} - 2w_{j,\nu} + w_{j-1,\nu}\right). \tag{6.26}$$

This is a FTCS scheme with the additional term

$$\frac{(\Delta x)^2}{2\Delta t}\,\delta_{xx} w_{j,\nu}\,,$$

representing the PDE

$$u_t + au_x = \zeta u_{xx} \quad \text{with } \zeta = \Delta x^2/2\Delta t\,.$$

That is, the stabilization is accomplished by adding artificial diffusion $\zeta u_{xx}$. The scheme (6.26) is said to have *numerical dissipation*.

We return to the model problem (6.15) with $b > 0$. For the discretization of the $a\frac{\partial u}{\partial x}$ term we now apply the appropriate upwind scheme from (6.22) or (6.24), depending on the sign of the convection constant $a$. This noncentered first-order difference scheme can be written

$$\begin{aligned} w_{j,\nu+1} = w_{j,\nu} &- \gamma\tfrac{1-\mathrm{sign}(a)}{2}\left(w_{j+1,\nu} - w_{j,\nu}\right) \\ &- \gamma\tfrac{1+\mathrm{sign}(a)}{2}\left(w_{j,\nu} - w_{j-1,\nu}\right) \\ &+ \lambda(w_{j+1,\nu} - 2w_{j,\nu} + w_{j-1,\nu}) \end{aligned} \tag{6.27}$$

with parameters $\gamma, \lambda$ as defined in (6.18). For $a > 0$ the growth factors are

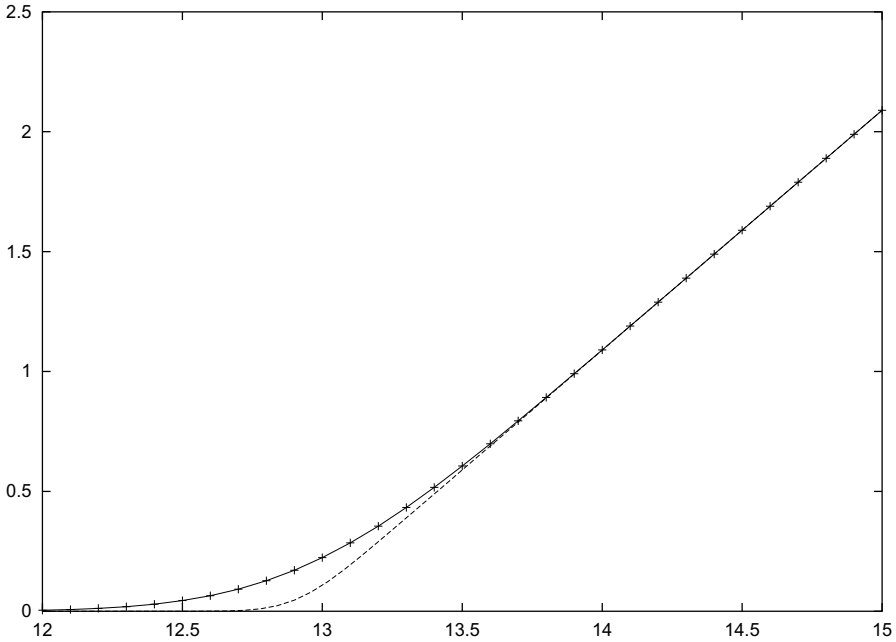


**Fig. 6.7.** European call, $K = 13$, $r = 0.15$, $\sigma = 0.01$, $T = 1$. Approximation $V(S, 0)$, calculated with upwind scheme for $\frac{\partial V}{\partial S}$ and $\Delta t = 0.01$, $\Delta S = 0.1$. Comparison with the exact Black–Scholes values (dashed)

$$G_k = 1 - \lambda(2 + \beta)(1 - \cos k\eta\Delta x) - \mathrm{i}\lambda\beta \sin k\eta\Delta x\,.$$

The analysis follows the lines of Section 6.4 and leads to the single stability criterion

$$\lambda \leq \frac{1}{2 + |\beta|}\,. \tag{6.28}$$

This inequality is valid for both signs of $a$ ($\longrightarrow$ Exercise 6.4). For $\lambda \ll \beta$ the inequality (6.28) is less restrictive than (6.20). For example, a hypothetical value of $\lambda = \frac{1}{50}$ leads to the bound $|\beta| \leq 10$ for the FTCS scheme (6.16) and to the bound $|\beta| \leq 48$ for the upwind scheme (6.27).

The Figures 6.7 and 6.8 show the Black–Scholes solution (dashed curve) and an approximation obtained by using the upwind scheme as in (6.27). No oscillations are visible, but the low order of the approximation can be seen from the moderate gradient, which does not reflect the steep gradient of the reality. The spurious wiggles have disappeared but the steep profile is heavily smeared. So the upwind scheme discussed above is a motivation to look for better methods (in Section 6.6).

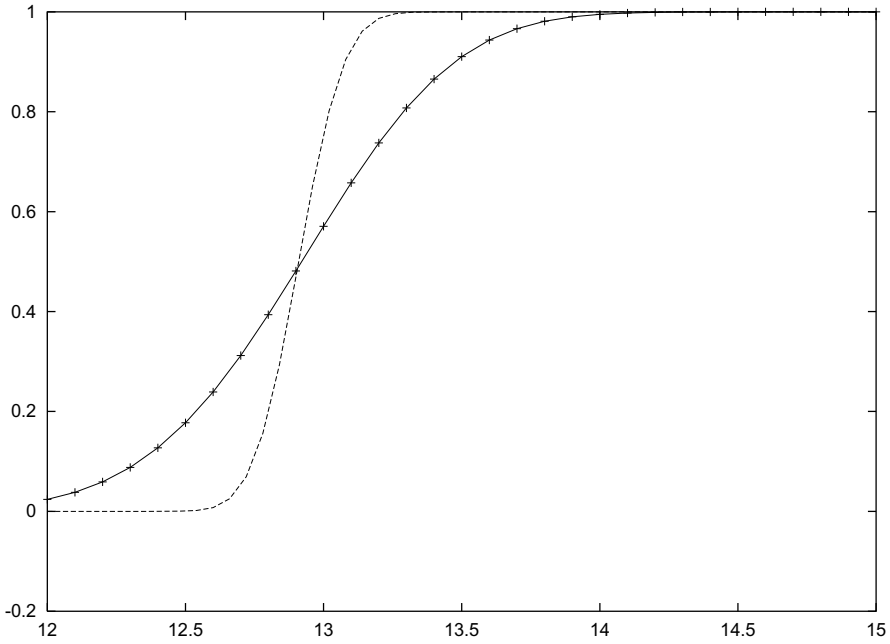


**Fig. 6.8.** Delta= $\frac{\partial V}{\partial S}(S,0)$, same data as in Figure 6.7

### 6.5.2 Dispersion

The spurious wiggles are attributed to *dispersion*. Dispersion is the phenomenon of different modes traveling at different speeds. We explain dispersion for the simple PDE $u_t + au_x = 0$. Consider for $t = 0$ an initial profile $u$ represented by a sum of Fourier modes, as in (6.17). Because of the linearity it is sufficient to study how the $k$th mode $\mathrm{e}^{\mathrm{i}kx}$ is conveyed for $t > 0$. The differential equation $u_t + au_x = 0$ conveys the mode without change, because $\mathrm{e}^{\mathrm{i}k[x-at]}$ is a solution. For an observer who travels with speed $a$ along the $x$-axis, the mode appears "frozen."

This does not hold for the numerical scheme. Here the amplitude and the phase of the $k$th mode may change. That is, the special initial profile of the Fourier mode

$$\mathrm{e}^{\mathrm{i}kx} = 1 \cdot \mathrm{e}^{\mathrm{i}k[x-0]}$$

changes to

$$c(t) \cdot \mathrm{e}^{\mathrm{i}k[x-d(t)]},$$

where $c(t)$ is the amplitude and $d(t)$ the phase (up to the traveler distance $at$). Their values must be compared to those of the exact solution.

To be specific, we study the upwind scheme for $u_t + au_x = 0$ $(a > 0)$,

$$\frac{w(x,t+\Delta t)-w(x,t)}{\Delta t}+a\frac{w(x,t)-w(x-\Delta x,t)}{\Delta x}=0\,.$$

Let $w(x,t)$ denote the exact solution for specified values of $\Delta x, \Delta t$. Apply Taylor's expansion to derive the *equivalent differential equation*

$$w_t + aw_x = \zeta w_{xx} + \xi w_{xxx} + O(\Delta^2)\,,$$

with the coefficients

$$\begin{aligned}\zeta &:= \frac{a}{2}(\Delta x - a\Delta t) &&= \frac{a}{2}\Delta x(1-\gamma)\,,\\ \xi &:= \frac{a}{6}(-\Delta x^2 + 3a\Delta t\Delta x - 2a^2\Delta t^2) &&= \frac{a}{6}\Delta x^2(1-\gamma)(2\gamma-1)\end{aligned}$$

depending on $\Delta x, \Delta t$. A solution can be obtained for the truncated PDE $w_t + aw_x = \zeta w_{xx} + \xi_{xxx}$. Substituting $w = \mathrm{e}^{\mathrm{i}(\omega t+kx)}$ with undetermined frequency $\omega$ gives $\omega$ and

$$w = \exp\{-\zeta k^2 t\}\cdot\exp\{\mathrm{i}k[x - t(\xi k^2 + a)]\}$$

as solution of the truncated PDE. This defines amplitudes $c(t)$ and phase shifts $d(t)$,

$$\begin{aligned}c_k(t) &= \exp\{-\zeta k^2 t\}\\ d_k(t) &= \xi k^2 t\,.\end{aligned}$$

The $w = c_k(t)\mathrm{e}^{\mathrm{i}k[x-at-d_k(t)]}$ represents the solution of the applied upwind scheme. It is compared to the exact solution $u = \mathrm{e}^{\mathrm{i}k[x-at]}$ of the model problem, for which all modes propagate with the same speed $a$ and without decay of the amplitude. The phase shift $d_k$ in $w$ due to a nonzero $\xi$ becomes more relevant if the wave number $k$ gets larger. That is, modes with different wave numbers drift across the finite-difference grid at different rates. Consequently, an initial signal represented by a sum of modes, changes its shape as it travels. The different propagation speeds of different modes $\mathrm{e}^{\mathrm{i}kx}$ give rise to oscillations. This phenomenon is called dispersion. (Note that in our scenario of the simple model problem with upwind scheme, for $\gamma = 1$ and $\gamma = \frac{1}{2}$ we have $\xi = 0$ and dispersion vanishes.)

A value of $|c(t)| < 1$ amounts to dissipation. If a high phase shift is compensated by heavy dissipation ($c \approx 0$), then the dispersion is damped and may be hardly noticeable.

For several numerical schemes, related values of $\zeta$ and $\xi$ have been investigated. For the influence of dispersion or dissipation see, for example, [Tho95], [QuSS00], [TaR00], [Str07]. Dispersion is to be expected for numerical schemes that operate on those versions of the Black–Scholes equation that have a convection term. This holds in particular for the $\theta$-methods as described in Section 4.6.1, and for the upwind scheme. Numerical schemes for the convection-free version $y_\tau = y_{xx}$ do not suffer from dispersion since $a = 0$.

## 6.6 High-Resolution Methods

The naive FTCS approach of the scheme (6.16) is only first-order in $t$-direction and suffers from severe stability restrictions. There are second-order approaches with better properties. A large class of schemes has been developed for so-called *conservation laws*, which in the one-dimensional situation are written

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0\,. \tag{6.29}$$

The function $f(u)$ represents the *flux* in the equation (6.29), which originally was tailored to applications in fluid dynamics. We introduce the method of Lax and Wendroff for the flux-conservative equation (6.29). Then we present basic ideas of high-resolution methods.

### 6.6.1 Lax–Wendroff Method

The Lax–Wendroff scheme is based on

$$u_{j,\nu+1} = u_{j,\nu} + \Delta t \frac{\partial u_{j,\nu}}{\partial t} + O(\Delta t^2) = u_{j,\nu} - \Delta t \frac{\partial f(u_{j,\nu})}{\partial x} + O(\Delta t^2)\,.$$

This expression makes use of (6.29) and replaces time derivatives by space derivatives. For suitably adapted indices the basic scheme is applied three times on a *staggered grid*. The staggered grid (see Figure 6.9) uses half steps of lengths $\frac{1}{2}\Delta x$ and $\frac{1}{2}\Delta t$ and intermediate mode numbers $j - \frac{1}{2}$, $j + \frac{1}{2}$, $\nu + \frac{1}{2}$. The main step is the second-order centered step (CTCS) with the center in the node $(j, \nu + \frac{1}{2})$ (square in Figure 6.9). This main step needs the flux function $f$ evaluated at approximations $w$ obtained for the two intermediate nodes $\left(j \pm \frac{1}{2}, \nu + \frac{1}{2}\right)$, which are marked by crosses in Figure 6.9. These two intermediate values are provided by the Lax–Friedrich steps (6.25).

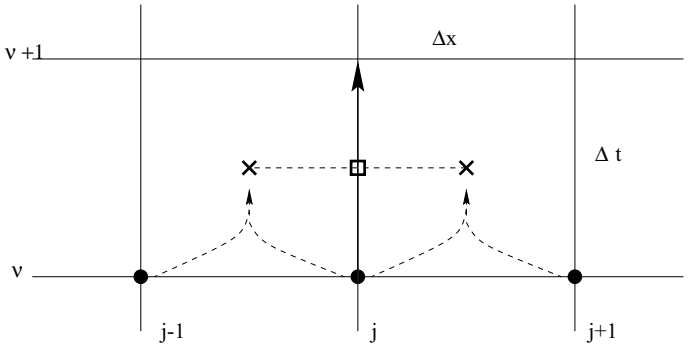


**Fig. 6.9.** Staggered grid for the Lax–Wendroff scheme.

**Algorithm 6.2 (Lax–Wendroff)**

$$\begin{aligned} w_{j+\frac{1}{2},\nu+\frac{1}{2}} &:= \tfrac{1}{2}(w_{j,\nu} + w_{j+1,\nu}) - \tfrac{\Delta t}{2\Delta x}\,(f(w_{j+1,\nu}) - f(w_{j,\nu})) \\ w_{j-\frac{1}{2},\nu+\frac{1}{2}} &:= \tfrac{1}{2}(w_{j-1,\nu} + w_{j,\nu}) - \tfrac{\Delta t}{2\Delta x}\,(f(w_{j,\nu}) - f(w_{j-1,\nu})) \\ w_{j,\nu+1} &:= w_{j,\nu} - \tfrac{\Delta t}{\Delta x}\left(f(w_{j+\frac{1}{2},\nu+\frac{1}{2}}) - f(w_{j-\frac{1}{2},\nu+\frac{1}{2}})\right) \end{aligned} \tag{6.30}$$

The half-step values $w_{j+\frac{1}{2},\nu+\frac{1}{2}}$ and $w_{j-\frac{1}{2},\nu+\frac{1}{2}}$ are provisional and discarded after $w_{j,\nu+1}$ is calculated. A stability analysis for the special case $f(u) = au$ in equation (6.29) (that is, of equation (6.21)) leads to the CFL condition as before. The Lax–Wendroff step is centered and of second order in both $x$ and $t$. This explicit method fits well discontinuities and steep fronts as the Black–Scholes delta-profile in Figures 6.5 and 6.8. But there are still spurious wiggles in the vicinity of steep gradients. The Lax–Wendroff scheme produces oscillations near sharp fronts. We need to find a way to damp out the oscillations.

### 6.6.2 Total Variation Diminishing

Since $u_t + au_x$ convects an initial profile $F(x)$ with velocity $a$, a monotonicity of $F$ will be preserved for all $t > 0$. So it makes sense to require also a numerical scheme to be *monotonicity preserving*. That is,

$$\begin{aligned} w_{j,0} \le w_{j+1,0} \;\text{ for all }\; j \quad &\Rightarrow \quad w_{j,\nu} \le w_{j+1,\nu} \;\text{ for all }\; j,\nu \ge 1 \\ w_{j,0} \ge w_{j+1,0} \;\text{ for all }\; j \quad &\Rightarrow \quad w_{j,\nu} \ge w_{j+1,\nu} \;\text{ for all }\; j,\nu \ge 1\,. \end{aligned}$$

A stronger requirement is that oscillations be diminished. To this end we define the *total variation* of the approximation vector $w^{(\nu)}$ at the $\nu$-th time level as

$$\mathrm{TV}(w^{(\nu)}) := \sum_j |w_{j+1,\nu} - w_{j,\nu}|\,. \tag{6.31}$$

The aim is to construct a method that is *total variation diminishing* (TVD),

$$\mathrm{TV}(w^{(\nu+1)}) \le \mathrm{TV}(w^{(\nu)}) \;\text{ for all }\; \nu\,.$$

Before we come to a criterion for TVD, note that the schemes discussed in this section are explicit and of the form

$$w_{j,\nu+1} = \sum_l d_l\, w_{j+l,\nu}\,. \tag{6.32}$$

For example, the upwind scheme (6.22) for $a > 0$

$$w_{j,\nu+1} = (1-\gamma)w_{j,\nu} + \gamma w_{j-1,\nu}$$

has two coefficients in (6.32), $d_{-1} = \gamma$ and $d_0 = 1 - \gamma$. The coefficients $d_l$ decide whether the scheme (6.32) is monotonicity preserving or TVD.

**Lemma 6.3 (monotonicity and TVD)**

(a) The scheme (6.32) is monotonicity preserving if and only if $d_l \geq 0$ for all $d_l$.

(b) The scheme (6.32) is total variation diminishing (TVD) if and only if

$$d_l \geq 0 \ \text{ for all } \ d_l\,, \ \text{ and } \ \sum_l d_l \leq 1\,.$$

The proof of (a) is left to the reader; for proving (b) the reader may find help in [Wes01], see also [Krö97]. As a consequence of Lemma 6.3 note that TVD implies monotonicity preservation.

The criterion of Lemma 6.3 is straightforward to check. For example, we can be certain now about the upwind scheme's monotonicity preservation shown in Figures 6.7, 6.8. The Lax–Wendroff scheme satisfies $d_l \geq 0$ for all $l$ only in the exceptional case $\gamma = 1$. For practical purposes, in view of nonconstant coefficients $a$, the Lax–Wendroff scheme is not TVD. For $f(u) = au$, the upwind scheme (6.22) and the Lax–Friedrichs scheme (6.25) are TVD for $|\gamma| \leq 1$ ($\longrightarrow$ Exercise 6.6).

### 6.6.3 Numerical Dissipation

For clarity we continue to discuss the matters for the linear scalar equation (6.21),

$$u_t + au_x = 0\ , \ \text{for } a > 0\,.$$

For this equation it is easy to substitute the two provisional half-step values of the Lax–Wendroff algorithm into the equation for $w_{j,\nu+1}$. Then a straightforward calculation shows that the Lax–Wendroff scheme can be obtained by adding a diffusion term to the upwind scheme (6.22). To show this, make use of the difference operator

$$\delta_x^- w_{j,\nu} := w_{j,\nu} - w_{j-1,\nu} \tag{6.33}$$

and rewrite the upwind scheme as

$$w_{j,\nu+1} = w_{j,\nu} - \gamma\delta_x^- w_{j,\nu}\ , \quad \gamma = \frac{a\,\Delta t}{\Delta x}\,.$$

The reader may check that the Lax–Wendroff scheme is obtained by adding the term

$$-\delta_x^-\{\tfrac{1}{2}\gamma(1-\gamma)(w_{j+1,\nu} - w_{j,\nu})\} \tag{6.34}$$

to the upwind scheme. So the Lax–Wendroff scheme is rewritten

$$w_{j,\nu+1} = w_{j,\nu} - \gamma\delta_x^- w_{j,\nu} - \delta_x^-\{\tfrac{1}{2}\gamma(1-\gamma)(w_{j+1,\nu} - w_{j,\nu})\}\,.$$

That is, the Lax–Wendroff scheme is the first-order upwind scheme plus the term (6.34), which is

$$-\tfrac{1}{2}\gamma(1-\gamma)(w_{j+1,\nu} - 2w_{j,\nu} + w_{j-1,\nu})\,.$$

Hence the added term is —similar as for the Lax–Friedrichs scheme (6.26)— the discretized analogue of the artificial diffusion

$$-\tfrac{1}{2}a\Delta t(\Delta x - a\Delta t)u_{xx}\,.$$

Adding this artificial dissipation term (6.34) to the upwind scheme makes the scheme a second-order method.

The aim is to find a scheme that will give us neither the wiggles of the Lax–Wendroff scheme nor the smearing and low accuracy of the upwind scheme. On the other hand, we wish to benefit both from the second-order accuracy of the Lax–Wendroff scheme and from the smoothing capabilities of the upwind scheme. A core idea is not to add the same amount of dissipation everywhere along the $x$-axis, but to add artificial dissipation in the right amount where it is needed. This flexibility is achieved by a proper factor on the diffusion (6.34). The resulting hybrid scheme will be of Lax–Wendroff type when the gradient is flat, and will be upwind-like at strong gradients of the solution. The decision on how much dissipation to add will be based on the solution.

In order to meet the goals, high-resolution methods control the artificial dissipation by introducing a *limiter* $\ell_{j,\nu}$ such that

$$w_{j,\nu+1} = w_{j,\nu} - \gamma\delta_x^- w_{j,\nu} - \delta_x^-\{\,\ell_{j,\nu}\,\tfrac{1}{2}\gamma(1-\gamma)(w_{j+1,\nu} - w_{j,\nu})\}\,. \tag{6.35}$$

Obviously this hybrid scheme specializes to the upwind scheme for $\ell_{j,\nu} = 0$ and is identical to the Lax–Wendroff scheme for $\ell_{j,\nu} = 1$. Accordingly, $\ell_{j,\nu} = 0$ should be chosen for strong gradients in the solution profile and $\ell_{j,\nu} = 1$ for smooth sections. To check the smoothness of the solution one defines the *smoothness parameter*

$$q_{j,\nu} := \frac{w_{j,\nu} - w_{j-1,\nu}}{w_{j+1,\nu} - w_{j,\nu}}\,. \tag{6.36}$$

The limiter $\ell_{j,\nu}$ will be a function of $q_{j,\nu}$. We now drop the indices $j, \nu$. For $q \approx 1$ the solution will be considered smooth, so we require the function $\ell = \ell(q)$ to satisfy $\ell(1) = 1$ to reproduce the Lax–Wendroff scheme. Several strategies have been suggested to choose the limiter function $\ell(q)$ such that the scheme (6.35) is total variation diminishing. For a thorough discussion of this matter we refer to [Swe84], [Krö97], [Tho99]. One example of a limiter function is the van Leer limiter, which is defined by

$$\ell(q) = \begin{cases} 0 & ,\ q \le 0 \\ \frac{2q}{1+q} & ,\ q > 0 \end{cases} \tag{6.37}$$

The above principles of high-resolution methods have been applied successfully to financial engineering. The transfer of ideas from the simple problem (6.21) to the Black–Scholes world is quite involved. The methods are

TVD for the Black–Scholes equation, which is in nonconservative form. Further the methods can be applied to nonuniform grids, and to implicit methods. The application of the Crank-Nicolson approach can be recommended. The equations (6.36), (6.37) introduce a nonlinearity in $w^{(\nu+1)}$. Hence nonlinear equations are solved for each time step $\nu$; Newton's method is applied to calculate the approximation $w^{(\nu+1)}$ [ZvFV98].

## 6.7 Penalty Method for American Options

As we have seen in Chapter 4, the PDE description of an American-style option leads to a linear complementarity problem (LCP), which was restated in Problem 4.12 as an equation under an inequality as side condition. Such problems can be solved numerically by imposing a penalty in case the inequality is violated. For motivation see Section 4.5.4, and study the simple setting of Exercise 6.8. Penalty methods have been applied repeatedly for the pricing of American options, see for instance [FoV02], [NiST02], [KoLM07]. Here we describe the approach of [NiST08].

### 6.7.1 LCP Formulation

Similar as in Section 4.5.3 we denote the $n$-dimensional Black–Scholes operator of (6.2b)

$$\mathcal{L}_{\mathrm{BS}}(V) := \frac{1}{2}\sum_{i,j=1}^{n} \rho_{ij}\,\sigma_i\sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j} + \sum_{i=1}^{n}(r-\delta_i)S_i\frac{\partial V}{\partial S_i} - rV \tag{6.38}$$

and the payoff by $\Psi(S_1,\ldots,S_n)$. For example, for a basket put,

$$\Psi(S_1,\ldots,S_n) = \left(K - \sum_{i=1}^{n} c_i S_i\right)^+ .$$

With the vector $S := (S_1,\ldots,S_n)$ the LCP is

$$\begin{aligned} &(V-\Psi)\left(\frac{\partial V}{\partial t} + \mathcal{L}_{\mathrm{BS}}(V)\right) = 0 \\ &\frac{\partial V}{\partial t} + \mathcal{L}_{\mathrm{BS}}(V) \le 0 \\ &V \ge \Psi \end{aligned} \tag{6.39}$$

In addition, the terminal condition $V(S,T) = \Psi(S)$ must hold, and boundary conditions. Since the domain is $S > 0$, there are $n$ bounding planes given by $S_i = 0$, $i = 1,\ldots,n$. For each $i$ let

$$\mathcal{D}_i := \{\,(S_1,\ldots,S_{i-1},0,S_{i+1},\ldots,S_n) \mid S_j > 0 \text{ for } j \ne i\,\}$$

denote the domain of the associated $(n-1)$-dimensional American option problem with the same terms, and $G_i(S,t)$ for $S \in \mathcal{D}_i$ be its solution. Then the boundary conditions for the bounding planes $S_i = 0$ are defined by

$$V(S,t) = G_i(S,t) \quad \text{for } S \in \mathcal{D}_i \tag{6.40}$$

for all $i = 1, \ldots, n$. Note that these boundary conditions amount to the recursive solution of all lower-dimensional American option problems. This is an enormous amount of work for larger $n$, and limits the approach to small values of the dimension. The final item to be specified are the boundary conditions for $S_i \to \infty$. For the case of a put,

$$\lim_{S_i \to \infty} V(S,t) = 0 \quad \text{for all } i\,.$$

The above equations define the LCP for an $n$-asset American option under the Black–Scholes model.

### 6.7.2 Penalty Formulation

In the following, we stay with the American put with a basket payoff. For a penalty approach, replace the LCP formulation (6.39) by

$$\begin{aligned} &\frac{\partial V^{\epsilon,C}}{\partial t} + \mathcal{L}_{\mathrm{BS}}(V^{\epsilon,C}) + \frac{\epsilon\, C}{V^{\epsilon,C} + \epsilon - q} = 0 \\ &\text{with } q := K - \sum_{i=1}^{n} c_i S_i\,. \end{aligned} \tag{6.41}$$

$q$ is the basic part of the basket's payoff. We call the solution of the penalty formulation (6.41) $V^{\epsilon,C}$; it is supposed to approximate $V$. Clearly, the value function $V$ and its approximation $V^{\epsilon,C}$ should both satisfy $V \geq q$. The parameter $\epsilon$ in the penalty term

$$p := \frac{\epsilon\, C}{V^{\epsilon,C} + \epsilon - q} \tag{6.42}$$

must be chosen small with $0 < \epsilon \ll 1$. The parameter $C > 0$ is a tune factor to be fixed later. For $V^{\epsilon,C} \gg q$, the penalty term is of the order $\epsilon$, and (6.41) approximates the Black–Scholes equation. As $V^{\epsilon,C}$ approaches the payoff, $V^{\epsilon,C} \approx q$, the penalty term $p$ approaches the value $C > 0$, and

$$\frac{\partial V^{\epsilon,C}}{\partial t} + \mathcal{L}_{\mathrm{BS}}(V^{\epsilon,C}) \approx -C < 0\,.$$

This reflects the complementarity of American options. Note that the equation (6.41) is *nonlinear* in $V$.[7]

[7] Actually, the LCP (6.39) is nonlinear as well, which is not correctly reflected by the name "LCP".

### 6.7.3 Discretization of the Two-Factor Model

For the discretization of the American-style basket put we restrict ourselves to the case $n = 2$. Then the lower-dimensional American put problems are the plain-vanilla cases discussed in Chapter 4, and the corresponding standard value functions $G_1(S_2, t)$ for $S_1 = 0$ and $G_2(S_1, t)$ for $S_2 = 0$ can be considered "known" or delegated to a subalgorithm. The functions $G_1$ and $G_2$ are defined by the Black–Scholes equation/inequality, and by their payoff and volatility:

$$\begin{aligned} &G_1(S_2, t) \text{ with payoff } (K - c_2 S_2)^+, \text{ volatility } \sigma_2\,, \\ &G_2(S_1, t) \text{ with payoff } (K - c_1 S_1)^+, \text{ volatility } \sigma_1\,. \end{aligned}$$

Here we apply a standard finite-difference scheme, widely analogous as in Chapter 4. The nonlinearity of the PDE (6.41) prevents a transformation such as (4.3). Hence the discretization is applied to (6.41) directly. For ease of notation, we use the variables

$$x := S_1,\ y := S_2\,,$$

and $\rho$ for $\rho_{12}$. Then the penalty problem (6.41) for $V^{\epsilon,C}(x, y, t)$ is restated as (the superscript $\epsilon, C$ of $V^{\epsilon,C}$ is dropped)

$$\begin{aligned} &\frac{\partial V}{\partial t} + \frac{1}{2}\sigma_1^2 x^2 \frac{\partial^2 V}{\partial x^2} + \frac{1}{2}\sigma_2^2 y^2 \frac{\partial^2 V}{\partial y^2} + \rho\sigma_1\sigma_2 xy \frac{\partial^2 V}{\partial x \partial y} \\ &+ (r - \delta_1) x \frac{\partial V}{\partial x} + (r - \delta_2) y \frac{\partial V}{\partial y} - rV + \frac{\epsilon\, C}{V + \epsilon - q} = 0 \end{aligned} \tag{6.43}$$

with terminal and boundary conditions. For a put with basket payoff these are:

$$\begin{aligned} &q(x, y) := K - c_1 x - c_2 y \\ &\Psi(x, y) := (q(x, y))^+ \\ &V^{\epsilon,C}(x, y, T) = \Psi(x, y) \\ &V^{\epsilon,C}(x, 0, t) = G_2(x, t) \\ &V^{\epsilon,C}(0, y, t) = G_1(y, t) \\ &\lim_{x\to\infty} V^{\epsilon,C}(x, y, t) = \lim_{y\to\infty} V^{\epsilon,C}(x, y, t) = 0\,, \end{aligned}$$

for $0 \le t \le T$, $x \ge 0$, $y \ge 0$. An equidistant grid on the truncated domain

$$0 \le x \le x_{\max},\ 0 \le y \le y_{\max},\ 0 \le t \le T$$

is defined by $i_{\max}$, $j_{\max}$ and $\nu_{\max}$ subintervals,

$$\Delta x := \frac{x_{\max}}{i_{\max}},\ x_i := i\Delta x,\ i = 0, \ldots, i_{\max}$$
$$\Delta y := \frac{y_{\max}}{j_{\max}},\ y_j := j\Delta y,\ j = 0, \ldots, j_{\max}$$
$$\Delta t := \frac{T}{\nu_{\max}},\ t_\nu := \nu\Delta t,\ \nu = \nu_{\max}, \ldots, 0\,.$$

Furthermore, we use the notations

$$q_{i,j} := q(x_i, y_j)\,,$$
$$w^\nu_{i,j} \text{ approximation to } V^{\epsilon,C}(x_i, y_j, t_\nu)\,.$$

To simplify the exposition, we choose $i_{\max} = j_{\max}$, $x_{\max} = y_{\max}$ and use the notation $h := \Delta x = \Delta y$. The difference quotients are defined in Chapter 4, except for the mixed second-order derivative, which is discretized by the second-order term

$$\delta_{xy} w^\nu_{i,j} := \frac{1}{2h^2}(w^\nu_{i+1,j+1} - w^\nu_{i+1,j} - w^\nu_{i,j+1} + 2w^\nu_{i,j} - w^\nu_{i-1,j} - w^\nu_{i,j-1} + w^\nu_{i-1,j-1})$$

By stability reasons ($\longrightarrow$ Section 6.4, 6.5) the first-order derivatives with respect to $x$ and $y$ are discretized by upwind schemes. For $\delta_1 \le r$, $\delta_2 \le r$, the upwind schemes are

$$\delta_x w^\nu_{i,j} := \frac{w^\nu_{i+1,j} - w^\nu_{i,j}}{h}\,,$$
$$\delta_y w^\nu_{i,j} := \frac{w^\nu_{i,j+1} - w^\nu_{i,j}}{h}\,,$$

since the integration is backward in time. Substituting all difference quotients into (6.43) is routine.

As in Chapter 4, we may choose among explicit or implicit schemes. The difference quotient

$$\delta_t w^\nu_{i,j} := \frac{w^{\nu+1}_{i,j} - w^\nu_{i,j}}{\Delta t}$$

for the time derivative $\frac{\partial V}{\partial t}$ leads to an explicit scheme when the difference quotients with respect to $x, y$ are evaluated at level $\nu + 1$, and leads to an implicit scheme when the evaluation is at level $\nu$. In the latter case, since we integrate backward in time, $w^{\nu+1}$ is considered as calculated and the $w^\nu$ are to be calculated next. For the explicit scheme, stability requirements lead to severe restrictions on the step size $\Delta t$, and to a slow algorithm; it will not be discussed further.

But for the implicit scheme, the nonlinear penalty term (6.42) makes a difference. In case we plug in $w^\nu_{i,j}$ for $V^{\epsilon,C}$, the equation to be solved at time level $t_\nu$ is nonlinear and requires an iterative solution. To speed up a Newton iteration, good initial guesses must be made available. These are given by the previous time level, provided the time steps $\Delta t$ are small. Such a restriction on $\Delta t$ *due to the nonlinearity* may make the method expensive. But there

is an alternative. When $w_{i,j}^{\nu+1}$ is used for $V^{\epsilon,C}$ in the penalty term, then the nonlinearity at time level $t_\nu$ is known, and for each $\nu$ only a linear system needs to be solved. This procedure is called *semi-implicit* or linear-implicit. The alternative of a fully nonlinear equation [with $w_{i,j}^{\nu}$ in (6.42)] is referred to as fully implicit.

The semi-implicit scheme now reads

$$\frac{w_{i,j}^{\nu+1} - w_{i,j}^{\nu}}{\Delta t} + \frac{1}{2}\sigma_1^2 x_i^2\, \delta_{xx} w_{i,j}^{\nu} + \frac{1}{2}\sigma_2^2 y_j^2\, \delta_{yy} w_{i,j}^{\nu} + \rho\sigma_1\sigma_2 x_i y_j\, \delta_{xy} w_{i,j}^{\nu}$$
$$+ (r-\delta_1) x_i\, \delta_x w_{i,j}^{\nu} + (r-\delta_2) y_j\, \delta_y w_{i,j}^{\nu} - r w_{i,j}^{\nu} + \frac{\epsilon\, C}{w_{i,j}^{\nu+1} + \epsilon - q_{i,j}} = 0$$

for $\nu = \nu_{\max} - 1, \ldots, 0$, and $w_{i,j}^{\nu_{\max}} = \Psi(x_i, y_j)$. We leave it to the reader to plug in the difference quotients, to organize the equation, and to introduce a matrix-vector notation for the equation to be solved at time level $t_\nu$.

In [NiST08] the explicit, the semi-implicit, and the fully implicit schemes were analyzed for the uncorrelated case $\delta = 0$. In these numerical experiments it turned out that the semi-implicit variant is recommendable in terms of accuracy and costs. In case

$$C \geq rK\,,\ \Delta t \leq \frac{\epsilon}{rK} \tag{6.44}$$

holds, the semi-implicit method satisfies the required inequality

$$w_{i,j}^{\nu} \geq \Psi(x_i, y_j)$$

for all $\nu$, see [NiST08]. This restricts the step size $\Delta t$ to a small value. Hence, one will not choose a too small value of $\epsilon$ and do without high demands on the accuracy of $V^{\epsilon,C}$. For example, one chooses $\epsilon = 0.01$ or $\epsilon = 0.001$. But for the fully implicit method the step size $\Delta t$ must be restricted too in order to maintain the convergence of the Newton method. And the mild bound on $\Delta t$ in (6.44) does not depend on $h$ (as would do the bound of the explicit method). Our experiments indicate an $O(\epsilon)$ error of $V^{\epsilon,C}$.

## Notes and Comments

*on Section 6.1:*

For barrier options we refer, for example, to [ZvFV99], [StWH99], [Ave00], [PoFVS00], [ZvVF00]. [DaL10] suggest a tree method with an initial trinomial step tuned so that the following tree has layers coinciding with the barrier. For lookback options we mention [Kat95], [FoVZ99], [Dai00]. [Haug98] is a rich source of analytical formula for option pricing.

*on Section 6.2:*

To see how the multidimensional volatilities of the model enter into a lumped volatility, consult [Shr04]. Other multidimensional PDEs arise when stochastic volatilities are modeled with SDEs, see [BaR94], [ZvFV98a], [Oos03], [HiMS05], [HaH10], or Example 5.7. A list of exotic options with various payoffs is presented in Section 19.2 of [Deu01]. Also the $n$-dimensional PDEs can be transformed to simpler forms. This is shown for $n = 2$ and $n = 3$ in [Int07]. For the $n$-dimensional Black–Scholes problem, see [Kwok98], [AcP05], [CaD05]. An ADI method is applied to American options on two stocks in [ViZ02]. Refined ADI methods work with non-equidistant grids [HaH10]. Consult also the efficient operator splitting method [IkT09], which decouples the treatment of the early-exercise constraint and the solution of the linear system. Further higher-dimensional PDEs related to finance can be found in [TaR00].

*on Section 6.3:*

PDEs in the context of Asian options were introduced in [KeV90], [RoS95]. A reduction as in (6.8b) from $V(S, A, t)$ to $H(R, t)$ is called *similarity reduction*. The derivation of the boundary-value problem (6.12) follows [WiDH96]. For the discrete sampling discussed in Section 6.3.4 see [WiDH96], [ZvFV99]. The strategies introduced for Asian options work similarly for other path-dependent options. An overview on methods for Asian options, and a semi-analytical method are found in [Zha01].

*on Section 6.4:*

The von Neumann stability analysis is tailored to linear schemes and pure initial-value problems. It does not rigorously treat effects caused by boundary conditions. In this sense it provides a necessary stability condition for boundary-value problems. For a rigorous treatment of stability see [Tho95], [Tho99]. The stability analysis based on eigenvalues of iteration matrices as used in Chapter 4 is an alternative to the von Neumann analysis.

Spurious oscillations are special solutions of the difference equations and do not correspond to solutions of the differential equation. The spurious oscillations are not related to rounding errors. This may be studied analytically for the simple ODE model boundary-value problem $au' = bu''$, which is the steady state of (6.15), along with boundary conditions $u(0) = 0$, $u(1) = 1$. Here for mesh Péclet numbers $\frac{a\Delta x}{b} > 2$ the analytical solution of the discrete centered-space analog is oscillatory, whereas the solution $u(x)$ of the differential equation is monotone, see [Mor96]. The model problem is extensively studied in [PeT83], [Mor96]. The mesh Péclet number is also called "algebraic Reynold's number of the mesh."

*on Section 6.5:*

It is recommendable to derive the equivalent differential equation in Section 6.5.2.

*on Section 6.6:*

The Lax–Wendroff scheme is an example of a *finite-volume method*. Another second-order scheme for (6.21) is the *leapfrog* scheme $\delta_t^2 w + a\delta_x^2 w = 0$, which involves three time levels. The discussion of monotonicity is based on investigations of Godunov, see [Krö97], [Wes01]. The Lax–Wendroff scheme for (6.21) and $\gamma \geq 0$ can also be written

$$w_j^{\nu+1} = w_j^\nu - \tfrac{1}{2}\gamma(w_{j+1}^\nu - w_{j-1}^\nu) + \tfrac{1}{2}\gamma^2(w_{j+1}^\nu - 2w_j^\nu + w_{j-1}^\nu)\,.$$

(This version adopts the frequent notation $w_j^\nu$ for our $w_{j,\nu}$.) Here the diffusion term has a slightly different factor than (6.34). The numerical dissipation term is also called *artificial viscosity*. In [Wes01], p. 348, the Lax–Wendroff scheme is embedded in a family of schemes. A special choice of the family parameter yields a third-order scheme. The TVD criterion can be extended to implicit schemes and to schemes that involve more than two time levels. For the general analysis of numerical schemes for conservation laws (6.29) we refer to [Krö97].

*on Section 6.7:*

In [NiST08] the linear systems were solved iteratively with the bi-conjugate gradient method Bi-CGSTAB [vdV92], [Saad03]. Choosing $\Delta t$ small provides good initial guesses for the next time level, which accelerates the iteration. Hence the limitation $\Delta t \leq \frac{\epsilon}{rK}$ is not too severe in practice. In our experiments, the penalty method did not achieve better results than a simple binomial-tree method. For the convergence of penalty methods consult [FoV02]. A penalty method with a smooth penalty has been implemented with finite elements in [KoLM07]. The weak formulation (compare Section 5.4) works with the relatively simple choice of boundary conditions $V = \Psi$ along the boundary. Exercise 6.8 follows [NiST02].

*on other methods:*

Computational methods for exotic options are under rapid development. The universal binomial method can be adapted to exotic options [Kla01], [JiD04]. [TaR00] gives an overview on a class of PDE solvers. For barrier options see [ZvFV99], [ZvVF00], [FuST02]. For two-factor barrier options and their finite-element solution, see [PoFVS00]. PDEs for lookback options are given in [Bar97]. Using Monte Carlo for path-dependent options, considerable efficiency gains are possible with bridge techniques [RiW02], [RiW03]. For Lévy process models, see, for example, [ConT04], [AlO06]. We recommend to consult, for example, the issues of the *Journal of Computational Finance*.

## Exercises

**Exercise 6.1 Project: Monte Carlo Valuation of Exotic Options**

Perform Monte Carlo valuations of barrier options, basket options, and Asian options, each European style.

**Exercise 6.2 PDEs for Arithmetic Asian Options**

a) Use the higher-dimensional Itô-formula (→ Appendix B2) to show that the value function $V(S, A, t)$ of an Asian option satisfies

$$\mathrm{d}V = \left( \frac{\partial V}{\partial t} + S\frac{\partial V}{\partial A} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) \mathrm{d}t + \sigma S \frac{\partial V}{\partial S}\, \mathrm{d}W\,,$$

where $S$ is the price of the asset and $A$ its average.

b) Construct a suitable riskless portfolio and derive the Black–Scholes equation

$$\frac{\partial V}{\partial t} + S\frac{\partial V}{\partial A} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0\,.$$

c) Use the transformation $V(S, A, t) = \widetilde{V}(S, R, t) = SH(R, t)$, with $R = \frac{A}{S}$ and transform the Black–Scholes equation (6.5) to

$$\frac{\partial H}{\partial t} + \frac{1}{2}\sigma^2 R^2 \frac{\partial^2 H}{\partial R^2} + (1 - rR)\frac{\partial H}{\partial R} = 0\,.$$

d) From

$$R_{t+\,\mathrm{d}t} = R_t + dR_t\,, \quad \mathrm{d}S_t = \mu S_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t$$

derive the SDE

$$\mathrm{d}R_t = (1 + (\sigma^2 - \mu)R_t)\,\mathrm{d}t - \sigma R_t\,\mathrm{d}W_t$$

e) For

$$A_t := \frac{1}{t}\int_0^t S_\theta\,\mathrm{d}\theta$$

show $\mathrm{d}A = \frac{1}{t}(S - A)\,\mathrm{d}t$ and derive the PDE

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} + \frac{1}{t}(S - A)\frac{\partial V}{\partial A} - rV = 0\,.$$

**Exercise 6.3 Neumann Stability Analysis**

Assume a difference scheme in the form (6.32)

$$w_j^{(\nu+1)} = \sum_l d_l\, w_{j+l}^{(\nu)}$$

and make use of the Fourier transform (6.17)

$$w_j^{(\nu)} = \sum_{k=0}^{n-1} c_k^{(\nu)} \mathrm{e}^{\mathrm{i}k\eta j\Delta x} \quad \text{for } \eta = \frac{2\pi}{n\Delta x}\,.$$

a) What are the coefficients $d_l$ for the FTCS method (6.16)?

b) Prove linear independence

$$\sum_{k=0}^{n-1} \alpha_k \exp[\mathrm{i}\tfrac{2\pi}{n}kj] = 0 \quad \Longrightarrow \quad \alpha_k = 0 \ \text{ for all } k$$

*Hint:* FFT equivalence (C1.8).

c) Show

$$c_k^{(\nu+1)} = c_k^{(\nu)} \sum_l d_l \, \mathrm{e}^{\mathrm{i}k\eta l\Delta x}\,.$$

**Exercise 6.4 Upwind Scheme**

Apply von Neumann's stability analysis to

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = b\frac{\partial^2 u}{\partial x^2}\ ,\ a > 0,\, b > 0$$

using the upwind scheme for the left-hand side and the centered second-order difference quotient for the right-hand side.

**Exercise 6.5 Towards the Black–Scholes Equation**

a) For the model equation (6.2a) set up the vector $a$ and the matrix $b$ for the general notation (1.41).

b) Let $LL^{\mathrm{tr}}$ be the Cholesky decomposition of the $\rho$-matrix, and $\tilde{b} := bL$. Show

$$\mathrm{trace}(\tilde{b}\tilde{b}^{\mathrm{tr}} V_{SS}) = \sum_{i,j=1}^{n} \rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j}\,.$$

c) Show

$$\mathrm{d}V = \left[\frac{\partial V}{\partial t} + \sum_{i=1}^{n}(\mu_i - \delta_i)S_i\frac{\partial V}{\partial S_i} + \frac{1}{2}\sum_{i,j=1}^{n}\rho_{ij}\sigma_i\sigma_j S_i S_j \frac{\partial^2 V}{\partial S_i \partial S_j}\right]\mathrm{d}t + \sum_{i=1}^{n}\sigma_i S_i \frac{\partial V}{\partial S_i}\mathrm{d}W^{(i)}\,.$$

**Exercise 6.6 TVD of a Model Problem**

Analyze whether the upwind scheme (6.22), the Lax–Friedrichs scheme (6.25) and the Lax–Wendroff scheme (6.30) applied to the scalar partial differential equation

$$u_t + a u_x = 0\,, \quad a > 0,\ t \geq 0,\ x \in \mathbb{R}$$

satisfy the TVD property.
*Hint:* Apply Lemma 6.3.

**Exercise 6.7 Binomial Tree for Two Assets**

A two-asset binomial tree with $(x, y)$-coordinates representing the assets, and time-coordinate $t$, is assumed to develop as follows: Each node with position $(x, y)$ may develop for $t \to t + \Delta t$ with equal probabilities 0.25 to the four positions

$$(xu, yA),\ (xu, yB),\ (xd, yC),\ (xd, yD) \qquad (*)$$

for constants $u, d, A, B, C, D$.

a) Show that the tree is recombining for $AD = BC$.
*Hint:* Sketch the possible values in a $(x, y)$-plane.

Following [Rub94b], a tree is defined for interest rate $r$, asset parameters $\sigma_1, \sigma_2$, correlation $\rho$, and dividend rates $\delta_1, \delta_2$, by

$$\begin{aligned}
\mu_i &:= r - \delta_i - \sigma_i^2/2 \ \text{ for } i = 1, 2\\
u &:= \exp(\mu_1 \Delta t + \sigma_1 \sqrt{\Delta t})\\
d &:= \exp(\mu_1 \Delta t - \sigma_1 \sqrt{\Delta t})\\
A &:= \exp(\mu_2 \Delta t + \sigma_2 \sqrt{\Delta t}[\rho + \sqrt{1-\rho^2}])\\
B &:= \exp(\mu_2 \Delta t + \sigma_2 \sqrt{\Delta t}[\rho - \sqrt{1-\rho^2}])\\
C &:= \exp(\mu_2 \Delta t - \sigma_2 \sqrt{\Delta t}[\rho - \sqrt{1-\rho^2}])\\
D &:= \exp(\mu_2 \Delta t - \sigma_2 \sqrt{\Delta t}[\rho + \sqrt{1-\rho^2}])
\end{aligned}$$

For initial prices $x^0 := S_1^0$, $y^0 := S_2^0$, and time level $t_\nu := \nu \Delta t$, the $S_1$-components of the grid according to $(*)$ distribute in the same way as for the one-dimensional tree,

$$x_i^\nu := S_1^0 u^i d^{\nu - i} \ \text{ for } i = 0, \ldots, \nu\,.$$

b) Show that the second ($S_2$-)components belonging to $x_i^\nu$ are

$$y_{i,j}^\nu := S_2^0 \exp(\mu_2 \nu \Delta t)\, \exp\left(\sigma_2 \sqrt{\Delta t}\left[\rho(2i - \nu) + \sqrt{1-\rho^2}(2j - \nu)\right]\right)\,.$$

for $j = 0, \ldots, \nu$.
*Hints:* For $\nu \to \nu + 1$, $u$ corresponds to $i \to i + 1$, and $d$ corresponds to $i \to i$; $C \exp(2\sigma_2 \sqrt{\Delta t}) = B$.

c) Set up a computer program that implements this binomial method. Analogously as in Section 1.4 work in a backward recursion for $\nu = M, \ldots, 0$. For each time level $t_\nu$ set up the $(x, y)$-grid with the above rules and $\Delta t = T/M$. For $t_M = T$ fix $V$ by the payoff $\Psi$, and use for $\nu < M$

$$V_{i,j}^{\text{cont}} = \exp(-r\Delta t)\,\frac{1}{4}(V_{i,j}^{\nu+1} + V_{i+1,j}^{\nu+1} + V_{i,j+1}^{\nu+1} + V_{i+1,j+1}^{\nu+1})\,.$$

Test example: max call with $\Psi(S_1, S_2) = (\max(S_1, S_2) - K)^+$, $S_1^0 = S_2^0 = K = T = 1$, $r = 0.1$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$, $\rho = 0.25$, $\delta_1 = \delta_2 = 0$. For $M = 2000$ an approximation of the American-style option is 0.0309527, and for the European style 0.0164554.

**Exercise 6.8 Initial-Value Problem with Penalty Term**

Consider the ODE initial-value problem

$$u' = -u\,, \quad u(0) = 2$$

with the additional constraint

$$u(t) \geq 1\,.$$

a) Give an analytical solution.

b) Discuss for a value of $\epsilon$ with $0 < \epsilon \ll 1$ the initial-value problem

$$v' = -v + \frac{\epsilon}{v - 1 - \epsilon}\,, \quad v(0) = 2\,.$$

*Hint:* Do some numerical experiments.

c) Show that the solution $v(t)$ of the initial-value problem in b) satisfies

$$1 \leq v \leq 2\,, \quad v' \leq 0\,, \quad v'' \geq 0\,,$$

for $t \geq 0$.